Accepted Manuscript

Title: Incentives and children's dietary choices: a field experiment in primary schools

Author: Michèle Belot, Jonathan James, Patrick Nolen

PII: S0167-6296(16)30073-X
DOI: http://dx.doi.org/doi: 10.1016/j.jhealeco.2016.07.003
Reference: JHE 1954

To appear in: *Journal of Health Economics*

Received date: 16-1-2015
Revised date: 4-7-2016
Accepted date: 11-7-2016

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

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

*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 John Cawley, Joe Price, and two anonymous referees for their valuable comments. We also thank seminar and workshop participants at Aarhus, Bath, Bristol, Ghent, Glasgow, Edinburgh, Essex, European University Institute, Paris School of Economics, 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 Gender and Experiments Workshop in Alicante (2014), the European Economic Association annual conference (2013) and the European Society of Population Economics annual congress (2014), NBER Summer Institute (2015), and InstEAD Economics and Psychology Workshop. 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

July 4, 2016

Abstract

We conduct a field experiment in 31 primary schools in England to test the effectiveness of different temporary incentives on increasing choice and consumption of fruit and vegetables at lunchtime. In each treatment, pupils received a sticker for choosing a fruit or vegetable at lunch. They were eligible for an additional reward at the end of the week depending on the number of stickers accumulated, either individually (individual scheme) or in comparison to others (competition). Overall, we find no significant effect of the individual scheme, but positive effects of competition. For children who had margin to increase their consumption, competition increases choice of fruit and vegetables by 33% and consumption by 48%. These positive effects generally carry over to the week immediately following the treatment, but are not sustained effects six months later. We also find large differences in effectiveness across demographic characteristics such as age and gender.

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

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

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 mortality 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 incentivize 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)), lose weight (Cawley and Price (2013), Horwitz et al. (2013), Jeffery (2012)) and eat more fruit and vegetables (Just and Price (2013), Loewenstein et al. (2016)). 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 backgrounds. Both groups have been shown to have less healthy diets and are particularly resistant to change (see Belot and James (2011), 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.

We conduct a randomized field experiment in 31 primary schools across England and implement, for four weeks, two incentive schemes: an individual based incentive and a competitive incentive. Our sample includes classes in year 2 (pupils aged 6-7) and in year 5

[1]See Cawley (2015) 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)).

(pupils aged 10-11) to be able to investigate effects by age. In each treatment pupils were given a sticker for choosing or bringing in a fruit or vegetable at lunch. At the end of the week (Friday afternoon after lunch), each pupil had the opportunity to pick a larger prize. In the individual incentive scheme, if a pupil collected four stickers during the week she or he was allowed to choose a prize. 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. 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 after lunch at the end of the week so children would not engage in strategic behaviour.

Using incentives to encourage 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)). 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)). To the best of our knowledge, competitive incentives have not yet been studied in the consumption of fruit and vegetables in the context of nutrition. While this might have potential to increase the consumption of fruit and vegetables, it also has the threat of being effective only for boys or more competitive children 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 once incentives are removed.

We find that the competitive scheme works well overall, with no negative effects for any subgroup. The results of individual incentives are mixed, and the scheme has no overall effect. The competitive treatment is more effective for all demographic groups and, overall, is nearly three times as effective at getting children to consume a portion of fruit or vegetable 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, girls, pupils from poorer socioeconomic backgrounds, and younger children respond more positively to competition than to 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 socioeconomic backgrounds respond positively to the competitive treatment, although the estimated effect is not significantly different from the individual scheme. Using a competitive incentive could improve effectiveness by increasing the choice and consumption among 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, at least in the short term. 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 5 and conclude in Section 6.

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. They incentivized fruit and vegetable *consumption* at lunch over a span of two or three weeks. They compare the effectiveness of various individual incentive schemes (piece rate monetary payment, lottery, nickel - which were either immediate or delayed). While they find positive significant effects during the intervention period, they do not find evidence of medium run effects (they followed up for four weeks after the incentive was removed). In a follow-up study, Loewenstein et al. (2016) keep the incentive (a token with a value of 25 cents that could be redeemed at the school shop, school carnival, or book fair) constant but vary the length of time the incentives are in place (three or five weeks). They find the effect of the incentive persisted two months after it had been removed and the consumption rate was higher for the schools where the intervention lasted 5 weeks.

Our experiment has important design differences when compared to the two aforementioned studies. First, we incentivize *choice* of fruit and vegetables. Second, we compare individual and competitive schemes while they focused only on individual schemes. Third, we use a longer incentive period than Just and Price (2013). Fourth, we introduce a weekly prize that is relatively larger in value than our daily prizes. This means the incentive at the daily level is not independent of choices made on other days of the week. Finally, Loewenstein et al. (2016) did not have a control group, which, as we will see in our analysis, turned out to be important when estimating the longer term effects in our study; consumption of fruit and vegetables appears to follow an upward trend for our control group. Below, we will discuss our experimental design in detail and compare our findings to these two closely related studies.

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 the effects of poor nutrition and the benefits of healthy eating: individuals may overweight the initial costs of eating healthier and (or) underweight 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, younger children, and children from poorer socioeconomic backgrounds are more impatient than other children;[4] this could explain why children with those demographic characteristics are less likely to make healthy dietary choices. In that context, providing immediate incentives to eat healthily may prove to be a powerful tool to get these groups to respond.

Few studies compare the effectiveness of different interventions on changing diet behavior. List and Samek (2015) have looked at the effects of information only campaigns versus information campaigns coupled with individual based incentives and found the latter to be more effective. However, a number of studies in the weight-loss literature have used two or more treatment arms. These, for example, include comparing the use of individual, group, lottery-based, and deposit contracts (see Jeffery (2012), John et al.(2011), Kullgren et al. (2013), and Volpp et al. (2008)).

[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 socioeconomic backgrounds and Bettinger and Slonim (2007) for boys versus girls, and for older children versus younger ones.

There is a well-established literature showing that boys tend to be more competitive than girls (see Gneezy et. al (2003), Gneezy and Rustichini (2004), and Booth and Nolen (2012)) yet the effect of competitive incentives on health-related behaviors (and specifically nutrition) has yet to be examined. However, Kullgren et al. (2013) use an incentive scheme similar to our competitive one in a weight loss study. They do not label their scheme as competitive, they call it a group incentive, but it has similar features to our competition scheme. Participants were placed into groups of 5, the identities of the other 4 individuals were not revealed. The $500 incentive was split among participants in each group whose weight loss was below their monthly target.

The choice of subgroups to focus on is inspired by Belot and James (2011), who evaluated the effects of the Jamie Oliver "Feed Me Better Campaign" in England, which consisted of improving the nutritious quality of school meals. They evaluated the effects of the campaign on educational outcomes and found that boys and children from lower socioeconomic background responded less (or later) to the campaign. As mentioned in the introduction, there is a fair amount of evidence showing that these subgroups tend to respond less to health interventions.

Both of our treatments also relate to the idea of gamification, which is the introduction of game playing into non-game areas in order to make them more enjoyable and engaging. Hamari et al. (2014) review the empirical literature on gamification. Out of the twenty-four papers, though, they only review one that focuses on health.

3. Experimental Design

We conducted a field experiment in England to examine and compare the effectiveness of a competitive and individual based scheme. We recruited schools in a three step process. 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 at 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. Twelve 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.

A companion paper, Belot and James (2016), documents the selection process by which local authorities and schools chose to participate in this experiment. They find that, overall, out of 30 variables, the only significant difference between LEAs that participate in the experiment and those that do not is that those who participate have more schools. This suggests that our sample of LEAs is largely representative of the population. Table B1 has descriptive statistics of the LEAs that participate, LEAs that collaborate (responded to recruiting attempts) but do not have schools which eventually participate, and LEAs that did not collaborate (did not respond). It could be the case that LEAs with more schools participated because they found it easier to provide the names of at least three schools. Furthermore, to examine things at the school level, Table B2 compares schools in the experimental sample to schools that we approached but did not take part. There are no significant differences between schools that agreed to participate and those who did not.

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 year 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. Parents also filled out a survey and provided additional background information. We do not have personal, background, or choice and consumption information for the other pupils. Overall, 15.85% were not included in the analysis due to the absence of parental consent; we have data on 638 children.

Randomization

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 similar 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);[5] (iv) proportion of children eligible for free school meals; (v) proportion of children

[5] Since our treatment was assigned at the school level we needed to use the total number of pupils in years 2 and 5 in our randomization to make sure our sample size was roughly the same across treatment arms.

eating free school meals; (vi) per pupil expenditure; (vii) per pupil expenditure on catering; (viii) percent of children achieving level 4 in both English and Mathematics in the Key Stage 2 exams; (ix) average point scores of children in key stage 2 exams; (x) average percent of children absent on a given day; (xi) percent of children absent from Key Stage 2 exams; (xii) school type (religious or comprehensive); (xiii) whether a school was involved in the Food for Life Program; (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: socioeconomic background of the student body; school quality; student quality; and school type[6]. Using a random number generator, schools were assigned to one of the three treatment arms. We then conducted 45 hypothesis tests (control versus each treatment arm and between the treatment arms) to see if the control sample was different than either treatment arm or if the treatment arms were different based on the observable characteristics. If one of the hypothesis tests showed a significant difference at the 5% level we then re-started the randomization; we ran the randomization six times. This ensured that, *ex ante*, at the school level, our sample was balanced by treatment arm.

Treatments

The two treatments we designed incentivize *choice* (rather than *consumption*) of fruit or vegetables at lunch. We decided to incentivize 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, that is low-cost and simple to implement. Rewarding for choice removes any subjective judgement of the monitor to decide what constitutes an

[6]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. When pupils take their Key Stage 2 exams their performance is marked as achieving level 1-7. For pupils aged 11, they are expected to earn at least a level 4 on their math, science, and English exams. 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 program (http://www.foodforlife.org.uk) the aim of which is to enact a whole school food reform by teaching children about healthy eating; it focuses on the promotion of healthy eating and the value of sustainable food consumption. Variable (xiv) is the overall classification of the school based on its Office for Standards in Education, Childrenâ€™s Services and Skills (Ofsted) results. Ofsted regularly inspects schools and other service providers. Based on these inspections, schools are given an overall rating: 1 = outstanding; 2 = good; 3 = requires improvement; and 4 = inadequate. Variable (xv) is one aspect that is included in determining the overall Ofsted rating (it uses the same scale) and is based on the extent to which the pupils adopt a healthy lifestyle and are encouraged and enabled to eat and drink healthily.

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.[7] Finally, rewarding for choice rather than actually consuming an item negates the possibility of cheating. For example, if rewards were based on consumption, pupils may have an incentive to dispose of the fruit or vegetable, hide it, give it to a friend or try to mislead monitors regarding actual consumption. For these reasons, monitoring consumption is more likely to be reliable when choice is incentivized.

In both schemes, children received a sticker each day if they chose or brought in a fruit or vegetable at lunchtime.[8] 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.[9] 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 the following 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. 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 pupil could only collect a maximum of four stickers. The groups were changed each week so the children could not anticipate with whom they would be competing and, in this treatment as well, unused stickers did not carry over to the following week.

Timing

Before the interventions began, a background survey was sent to the parents that covered information on age, gender, ethnicity, primary

[7]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.

[8]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 to help at lunch time. They received a compensation for collecting the information for us.

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

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 recorded if a pupil chose a fruit or vegetable or brought a fruit or vegetable in with a packed lunch. And the monitors recorded whether the pupil had consumed none, some, or more than half of that fruit or vegetable. On Friday that week children took a food knowledge test and a "spot-the-difference" test.[10] 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; pupils weres asked to circle those 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 in the treatment schools were informed about the incentive scheme 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 incentivized to choose a fruit or vegetable for a period of four weeks. Each day a pupil chose or brought in a fruit or vegetable with a packed lunch 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 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 & Estimation Approach

In this section we set out our hypotheses and our empirical strategy.

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

4.1 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), Loewenstein et al. (2016), and List and Samek (2014). Furthermore, the effect is likely to differ by subgroups. Since boys, younger children, and children from poorer socioeconomic 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 campaigns (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 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.[11] By incentivizing children to take a fruit or vegetable our 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 vegetables they consumed; the rewards are only based on choice. As with choice, there is reason to expect that the effect on consumption will vary with gender, age, and socioeconomic background.

[11] 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.

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. There is indeed evidence that food preferences and eating habits form in childhood and track into adulthood.[12] 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 costs of switching to healthy behaviour (see Charness and Gneezy (2009) and Acland and Levy (2013)) could be effective. 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 fruit and vegetables in the week immediately following the intervention and six months later.

4.2 Estimation Approach

To test our hypotheses we estimate the following:

$$y_{it} = \alpha + \beta_1 \text{Weeks } 2-5_t + \beta_2(\text{Competition}_i \times \text{Weeks } 2-5_t) + \beta_3(\text{Individual}_i \times \text{Weeks } 2-5_t)$$
$$+\beta_4 \text{Week } 6_t + \beta_5(\text{Competition}_i \times \text{Week } 6_t) + \beta_6(\text{Individual}_i \times \text{Week } 6_t)$$
$$+ c_i + \varepsilon_{it}$$

where y_{it} equals 1 if pupil *i* chose a fruit or vegetable on day *t* and 0 otherwise. $\text{Weeks } 2-5_t$ is a dummy variable that equals 1 during weeks 2 to 5 (this is when the incentives were in place) and 0 otherwise. $\text{Week } 6_t$ equals 1 for each day during week 6, the period immediately after the incentives were removed, and 0 otherwise. We use two indicators to capture the effect of being in a school that was treated:

[12] See Kelder et. al. (1994), Resnicow et. al. (1998), and Singer et. al. (1995) for discussions.

Competition_i equals 1 if pupil i is in a school that was assigned to the competitive treatment and 0 otherwise; and Individual_i equals 1 if pupil i is in a school that was assigned to the individual incentive treatment. Therefore, parameters β_2 and β_3 capture the effect of the incentives when they were in place. The parameters β_5 and β_6 capture the effect of the treatments in the week immediately after the incentives were removed. Finally, c_i is an individual pupil fixed effect and ε_{it} is an error term.

Our primary estimation method is therefore a linear probability model (LPM)[13] with pupil fixed effects (FE). This 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.

The dependent variable in our regressions is bounded upwards (at 1); children who chose 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.

5. Results

5.1 Summary Statistics

We first compare our treatment and control schools in the baseline period. Table 1 Panel A 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-values 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

[13]An alternative approach would be to estimate a fixed effects logit or probit. However, given that our main estimation of interest is an interaction term and that there are problems with estimating the marginal effects for interaction terms in the probit and logit models (Ai and Norton (2003)) we use the LPM. Furthermore, the positive aspects of this approach as highlighted in Angrist and Piske (2009) (page 105-107).

looking at sub-samples, our analysis may contain less than 30 schools. Therefore, the standard clustering methods are not appropriate. To deal with this we correct for the potential clustering problems using 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.

Table 1 Panel A 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 at baseline.

Table 1 Panel B shows the summary statistics for the sample of pupils who, in the baseline week, chose a fruit or vegetable less than 100% of the time. This group is of interest because they are the ones who have some margin to increase their consumption of fruit and vegetables 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 at 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 chose a fruit or vegetable on a given day. Our measure consumption is a dummy variable equal to one if the pupil eats at least some of a fruit or vegetable on that day (we will refer to this variable as “try”).[14] Since the incentive was based on the total amount of healthy choices made in a week, we provide a descriptive overview of the weekly mean outcomes for choice and consumption in Figures 1 and 2.

[14]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.

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 vegetable 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 likelihood of choosing a fruit or vegetable increases among the control group, and catches up with 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. At 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 are more likely to choose a fruit or vegetable after returning from the mid-term break (denoted by the vertical line), and that this increase occurs faster for pupils in the treatment groups.

Figure 2 shows the effect of the treatments on trying a fruit or vegetable. In panel (a) we again see the full sample. At baseline there are no significant differences between the treatments and the control (refer to Table 1). The control group is much slower to increase their consumption of fruit or vegetables upon returning to school in comparison to choice; they only show a small increase in week three that seems to persist into 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% in week 1 compared to the baseline. 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 when looking at 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. Figure 2 panel (b) shows the share of pupils consuming fruit and vegetables in the individual and competitive schemes is similar at the start: around 70% in week 1 and 75% in week 2. In week 3 and 4, though, a gap emerges: the share in the individual scheme drops to around 68% while it stays around 75% in the competitive scheme. Overall, this figure shows that pupils are much less likely to increase 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.

The rise in the control group after the mid-term break is notable. One explanation is that monitoring alone can cause an increase in fruit and vegetable consumption. However we cannot, with certainty, attribute the trend observed in the control group to the monitoring itself, since we do not know what would have happened independently of monitoring. One would need data where students are unaware they are monitored, which we do not have and would likely be difficult to obtain. The trend is important as it suggests that other factors (besides our incentive schemes) are also having large impacts on consumption of fruit and vegetables. These other factors could be the monitoring itself, but could also be due to seasonal variations, and holiday interruptions.

5.3 Short and Medium Term Effects

We now turn to the average treatment effects for the main outcome variables of interest.

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. In all of our tables we report both the standard errors clustered at the school level using standard methods and the p-value from the wild bootstrap, as discussed previously, following Cameron, Gelbach, and Miller (2008). We find little effects of either incentive scheme on choice overall (Column [1]). The point estimates for competition and the individual 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)[15] we also find no significant effects: 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. During the baseline, children in year two were choosing a fruit or vegetable 85.2% of the time therefore the point estimate of -0.066 implies an 8% decrease in the probability

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

of choosing a fruit or vegetable. 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 they did in baseline.

Table 2A, presented in Appendix B, shows 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.

We also 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 than in the individual setting. These results carry over to the medium term as well. This suggests competition may be more effective at getting younger and poorer pupils to choose healthier items than an individual based incentive scheme.

In Table 3 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. Restricting the sample in this way reduces the sample to 215 pupils. Column [1] shows that the competition increased the probability of choosing a fruit or vegetable by 17.5 percentage points and we find evidence that the effect was sustained 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.

We next turn to examine the various subgroups. For each of these groups we have a reduced sample size. Of the 215 pupils we have 102 boys and 113 girls, 93 pupils are in year 2 and 122 are in year 5. We have 29 pupils who are eligible for free school meals and 179 who are not eligible, there are 7 whom we do not have information on their free school meal eligibility. We find that competition significantly increased the likelihood of consuming fruit or vegetables 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 boys responded positively to the incentive but we again find 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, 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 (see Appendix B).

When we compare the two treatments, looking at the results at the bottom of Table 3, we find that girls and younger pupils responded significantly more positively to the competition than then the individual incentive. These results suggest that competition is working well on incentivizing pupils who have room to increase their choice of healthier items at lunchtime. While, even for pupils with worse 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.[16] We do not condition the consumption variable or the regressions on whether a pupil chooses 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]). We find no

[16]We also monitored whether the children at more than half the portion they were served. We report these in Tables B3 and B4, the results are very similar to what we report for trying.

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, boys; 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 (Appendix B) 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 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* behavior[17]). 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

[17] We restrict the sample based on choice rather than trying in order to keep the samples consistent. We have estimated the effects on trying for the group who tried fruits and vegetables less than 100% in the baseline week and the results are qualitatively similar. These are available from the authors on request.

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 little 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 incentive 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 strong positive evidence that the competitive incentive encourages all subgroups to choose and consume more fruit and vegetables (Hypotheses 1 and 2) and that, for most groups, those effects persist even when the incentive is removed (Hypothesis 3). Furthermore boys and FSM pupils do respond positively to the competitive scheme (unlike under other interventions) while girls, 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 want to look at the costs of getting a pupil to try an additional fruit or vegetable under each scheme. Since our most robust results are from the intervention period and the week immediately following, we focus on the shorter term effectiveness. Furthermore we compare the results to a commonly used intervention to understand how each scheme compares to currently implemented programs.

At most, each pupil could win one prize and earn five stickers per week. The prizes were all under £1 (some were only £0.30) and the stickers were no more than £0.04 each. That means we spent, at most, £4.20 per pupil over the course of the intervention. Since we let all pupils participate in the experiment we will look at the effects on the overall population (we could look at cost effectiveness for only our group of interest, i.e. those from the less that 100% group, but it would likely be costly to identify them and we do not know if our results hold for a targeted intervention).

When looking at the individual incentive we find that, during the intervention, pupils increased the likelihood of trying a fruit or vegetable by 3.3 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 0.7 more fruit and vegetables because of the intervention or, that it cost, roughly, £6 to get a pupil to eat an additional fruit or vegetable.

The competition scheme was more effective than the individual scheme; it increased the likelihood that, for the overall sample, the probability of trying a fruit or vegetable increased by 11 percentage points during the intervention and by 6.7 percentage points immediately after the incentive was removed. Thus, for the first five weeks of our experiment pupils ate 2.5 additional fruit or vegetables. That means it cost £1.68, at most, to get a pupil to eat an additional fruit or vegetable under the competitive scheme.

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[18] and showed six minute videos[19] 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[20] implementing the program 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 fruit 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 fruit 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

[18] In Ireland, generally, there is no provision of food by schools. Pupils are expected to bring in a packed lunch.
[19] See http://www.fooddudes.co.uk for examples of the videos.
[20] See "Strategy for School Fruit Scheme" submitted by Ireland for the 2012/2013 school year.

program for 60,000 pupils at over €2 million; nearly double the costs we are considering when looking at just the food and prizes. 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 multifaceted individual incentive scheme that had to be augmented by videos, food provision, and teachers taking time to discuss the goals of the program.[21]

5.4 Choice and Consumption Dynamics

Having established that there are differences in the effectiveness of the incentive schemes we now move on to explain why the competitive scheme appears to work better in comparison to individual incentive scheme. In this section we will analyze 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 and examine what those differences may suggest.

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 if there is a discouragement effect we break the sample into two groups. Table 6 column [1] 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 [2] 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 [2]. However, for those who have missed the chance of getting a prize the effect of the individual incentive is estimated to be negative, though, it is insignificant. Therefore, the effect of the incentive wears off for those who miss the goal in the individual incentive scheme.

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 chose five fruit or vegetables, though, they were guaranteed a prize. Since children did not know who was in their group and

[21]While our `trying' variable does not equate to the actual eating of fruits and vegetables as examined by Horne et. al. (2009) our results on whether they eat "more than half" are more comparable. Those results predict the same cost effectiveness as looking at `trying' (refer to tables B3 and B4 in the appendix).

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.[22] Therefore, the choices made on Wednesday, Thursday, and Friday depend on each studentâ€™s subjective probability and may not differ by the consumption patterns earlier in the week. Indeed, this is what we see, the point estimates in columns [1] and [2] are not significantly different for the competitive treatment[23]. They are also not significantly different from the overall effect for the less than 100% group shown for the competitive treatment in Table 3, though, we cannot determine if the effect in Table 6 is different than zero[24].

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. In contrast, the probability of winning the prize in the competitive treatment depends on the subjective probability that the student holds, which is presumably always positive.

The effect of the competitive scheme on consuming at least part of a fruit or vegetable is similar to what we found for choice. Columns [3] and [4] show again a large positive effect of competition that is relatively independent of previous consumption patterns during the week (the point estimate for competition is not significantly different between columns [3] and [4])[25]. The individual incentive only has a significant effect for children who have not missed the chance to win a prize; the estimate in column [4] for the individual incentive is significantly different from that in column [3]. There is room for discouragement in the individual incentive scheme.

Summarizing, we find that each incentive scheme is associated with different dynamics of choice and consumption behavior. In particular, the incentive in the individual scheme disappears by the end of the week for some students. Thus the overall differences in choice and consumption could be due to the way the goals are defined, with, in particular, the known constant goal of the individual incentive causing discouragement.

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

[23] However, we can reject that the point estimates in columns [1] and [2] are the same for the individual incentive scheme.

[24] This could be due to the significant decrease in the sample size caused by the sample restrictions on the dynamic analysis.

[25] However, again, it is also not significantly different than the overall estimate for trying in Table 5.

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.[26]

To capture the long run effect, we now include two additional interaction terms in our main regressions. These are the two treatment indicators with a dummy variable denoting 6 months later. Table 7 and 8 present the results for choice and try respectively. To conserve space we only present the additional interaction terms. For both the overall sample and the restricted and for both choice and consumption we do not find evidence of any long run effect; when looking at the significance using the appropriate bootstrap method, no results are significant. We examine the same subgroups as we did in section 5.3 and none of the coefficients are precisely estimated. Therefore, there is no evidence for hypothesis 3 with regards to the longer term.

5.6 Food Knowledge and Other Outcomes

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, and 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 9 for the whole sample and in Table 10 for the sample of children

[26] 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.

who chose less than 100% in the first week. The effects across groups are not consistent and we fail to find evidence that the scores improved more in the treated schools. If anything, we find negative effects for the children in the individual incentive group (restricted sample). We only find a positive significant effect, see Table 10 column 7, for the Year 5 children in the competition treatment. Based on these results, we can safely say there is no evidence that the responses to the intervention are driven by learning.

An additional exercise we propose is to check whether the interventions affected other relevant outcomes that could partially explain the treatment effects we found[27]. 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 B5 and B6. Table B5 reports results for the whole sample, while Table B6 reports results for the less than 100% sample. We do not find any significant effect on attendance overall or by subgroup. 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 fruit or vegetables on offer at school. Table B7 and B8 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).

[27] We do not directly measure behavioral problems, or classroom disruptions. We did however run a questionnaire through head teachers after the intervention asking for feedback. We do not have any evidence (even anecdotal) that the incentive schemes affected pupils behaviour in the classroom.

6. Conclusion

This paper provides field evidence on how two incentive schemes 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 different effects and one cannot draw a unique conclusion on whether incentives can positively influence the likelihood that a student will eat healthily. 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. However, the individual incentive, in contrast, has very heterogeneous effects. Older children respond positively, while younger children are affected negatively. Second, when there were significant effects during the intervention period, we tend find that those effects carry over into the week after the incentives are removed. However, we find no evidence of behaviour change six months later; the effects of the temporary incentive appear to be short lived.

When looking at our individual incentive scheme we find smaller effects than those found in Just and Price (2013) and Loewenstein et al. (2016). Even though their intervention is similar in nature, there are important differences in the design that are worth mentioning. We reward choice instead of consumption, we reward students each day based on one day's consumption with a reward of smaller value (a sticker) compared to their daily reward. Furthermore, while they stick to rewards based on each day's consumption only, we reward subjects with a relatively larger valued prize at the end of the week based on choices throughout the week (which has the potential to cause discouragement as discussed in the section 5.4).

The subject pools are also different between our paper and both studies. The baseline consumption rate of fruit and vegetables was 33.2% in Just and Price (2013) and 39% in Loewenstein et al. (2016) while for our overall sample it is 76% and for our less than 100% sample it is still 46%. Therefore the differences in our results could highlight a non-linear effect of incentives based on the initial level of healthy eating. Or, since we have less people who can respond positively to the intervention, the upper bound of any effect we could estimate is likely lower.

Despite the design and sample differences, though, we do come to qualitatively similar results regarding the short term effects of individual incentives. However, we do not find evidence of habit formation as in Loewenstein et al. (2016), despite having an intervention of similar length.

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. It is even possible that some incentives 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

Ai, C. and Norton, E.C. (2003). 'Interaction terms in logit and probit models', *Economics Letters*, Volume 80, Issue 1, Pages 123-129

Angrist, J.D., and Pischke, J. (2008). 'Mostly harmless econometrics: An empiricist's companion', Princeton university press

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

Belot M., and James J. (2011). 'Healthy School Meals and Educational Outcomes', *Journal of Health Economics*, vol. 30(3), 489-504

Belot M., and James J. (2016). 'Partner Selection into Policy Relevant Field Experiments'. *Journal of Economic Behavior and Organization*, vol. 123, 31-56

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

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.

Cawley, J. (2015). 'An economy of scales: A selective review of obesity's economic causes, consequences and solutions'. *Journal of Health Economics*, vol. 43, pp. 244-268

Cawley, J., and Price, J.A., (2013). 'A case study of a workplace wellness program that offers financial incentives for weight loss.' *Journal of Health Economics*, vol. 32, pp. 794803.

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.

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.

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.

Hamari, J., Koivisto, J., and Sarsa, H. (2014). Does Gamification Work? - A Literature Review of Empirical Studies on gamification. In proceedings of the 47th Hawaii International Conference on System Sciences, Hawaii, USA, January 6-9, 2014.

Horne, P.J., Hardman, C.A., Lowe, C.F., Tapper, K., Le Noury, J., Madden, P., Patel, P., and Doody M. ' Increasing parental provision and children's consumption of lunchbox fruit and vegetables in Ireland: the Food Dudes intervention' *European Journal of Clinical Nutrition*, 63(5), 613-8

Horwitz, J.R., Kelly, B.D., Di Nardo, J.E., 2013. Wellness incentives in the workplace: cost savings through cost shifting to unhealthy workers. *Health Affairs*, 32, pp. 468476.

Jeffery, R. W. (2012). Financial Incentives and Weight Control. *Preventive Medicine*, 55(Suppl), S61S67.

John, L. K., Loewenstein, G., Volpp, K.G., (2011). "Empirical observations on longer-term use of incentives for weight loss", *Preventive Medicine*, Vol. 55, pp. S68-S74

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.

Kullgren, J.T., Troxel, A.B., Loewenstein, G., Asch, D.A., Norton, L.A., Wesby L., Yuanyuan Tao, M.S., Jingsan Zhu M.S., and Volpp, K.G. (2013), “Individual- Versus Group-Based Financial Incentives for Weight Loss: A Randomized, Controlled Trial”. *Ann Intern Med*. pp. 505-514.

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.

Loewenstein, G., Price, J., Volpp, K., (2016). ‘Habit formation in children: Evidence from incentives for healthy eating, Journal of Health Economics’, vol 45., pp. 47-54

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.

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.

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.

Robertson, 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.

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., John L.K., Troxel A.B., Norton L., Fassbender J., Loewenstein G. (2008). "Financial Incentive Based Approaches for Weight Loss: A Randomized Trial." *JAMA.* 300(22), pp.2631-2637.

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	*N*	Treatment: Individual	*N*	Treatment: Competition	*N*	C vs T1 p-value	C vs T2 p-value	T1 vs T2 p-value
	(C)		(T1)		(T2)				
Panel A: All Pupils									
Choice	0.841	1018	0.847	765	0.821	1014	0.925	0.769	0.713
Try	0.739	1056	0.769	644	0.72	1039	0.721	0.815	0.599
Eat more than half	0.554	1056	0.618	644	0.614	1039	0.352	0.571	0.985
Female	0.513	1018	0.438	765	0.558	1014	0.188	0.414	0.040
1st Language English	0.977	1018	0.983	746	0.931	983	0.945	0.244	0.152
White British	0.905	1014	0.926	747	0.805	982	0.771	0.322	0.254
Year 2	0.500	1018	0.537	765	0.619	1014	0.835	0.286	0.647

Free School Meal %	0.206	1009	0.197	736	0.154	982	0.901	0.406	0.515
School Dinner %	0.52	998	0.453	677	0.479	961	0.539	0.699	0.795
Packed Lunch %	0.479	998	0.547	677	0.521	961	0.531	0.671	0.795
Special dietary	0.053	1014	0.097	744	0.128	972	0.162	0.132	0.699
requirements %									
Specific health cond. %	0.144	1004	0.167	742	0.161	951	0.561	0.585	0.887
Ofsted overall score	2.066	1018	1.875	765	2.206	1014	0.418	0.569	0.244
Ofsted Health Score	1.396	1018	1.536	765	1.424	1014	0.633	0.971	0.667
Per pupil Expenditure	4097	1018	4126	765	3816	1014	0.941	0.370	0.280
Catering costs	112.1	1018	94.1	765	62.6	1014	0.573	0.236	0.336
Food for Life	0.21	1018	0.40	765	0.17	1014	0.364	0.815	0.292
Headcount girls	106	1018	122	765	123	1014	0.667	0.362	0.979
Headcount boys	114	1018	138	765	131	1014	0.625	0.358	0.875

Average point score	0.288	861	0.28	670	0.283	874	0.144	0.272	0.731
% Level 4 or >	0.815	861	0.789	670	0.751	874	0.607	0.200	0.571
in Eng/Maths									
Persistent Absence	0.024	907	0.017	726	0.021	874	0.671	0.831	0.693
Absence	0.054	907	0.051	726	0.054	874	0.569	0.959	0.677
Panel B: Restricted sample (Chose less than 100% Choice in baseline week)									
Choice	0.545	356	0.515	241	0.477	346	0.735	0.464	0.639
Try	0.455	343	0.458	225	0.375	365	0.977	0.388	0.300
Eat more than half	0.329	343	0.356	225	0.323	365	0.715	0.929	0.675
Female	0.396	356	0.419	241	0.575	346	0.769	0.064	0.084
1st Language English	0.961	356	0.965	231	0.946	333	0.889	0.777	0.659
White British	0.854	356	0.944	231	0.784	333	0.262	0.617	0.202
Year 2	0.382	356	0.303	241	0.624	346	0.771	0.048	0.348
Free School Meal %	0.154	351	0.102	226	0.162	333	0.635	0.947	0.533
School Dinner %	0.441	349	0.371	240	0.558	321	0.729	0.452	0.302
Packed	0.556	349	0.629	240	0.442	321	0.723	0.456	0.302

Lunch %									
Special dietary	0.028	356	0.108	231	0.177	328	0.104	0.072	0.350
requirements %									
Specific health cond. %	0.179	351	0.228	228	0.128	328	0.625	0.482	0.236
Ofsted overall score	2.169	356	2.079	241	2.263	346	0.613	0.759	0.422
Ofsted Health Score	1.346	356	1.485	241	1.468	346	0.815	0.749	0.965
Per pupil Expenditure	3727	356	3919	241	3743	346	0.282	1.009	0.521
Catering costs	84.2	356	77.1	241	40.5	346	0.823	0.112	0.188
Food for Life	0.24	356	0.06	241	0.12	346	0.545	0.667	0.675
Headcount girls	111	356	120	241	119	346	0.603	0.671	0.947
Headcount boys	116	356	133	241	128	346	0.434	0.595	0.773
Average point score	0.287	335	0.289	221	0.283	313	0.677	0.306	0.156
% Level 4 or >	0.838	335	0.827	221	0.752	313	0.813	0.152	0.138
in Eng/Maths									
Persistent	0.017	335	0.011	236	0.018	313	0.667	0.847	0.482

Absence									
Absence	0.052	335	0.047	236	0.053	313	0.539	0.915	0.490

All variables are evaluated for the first week, before the start of the treatment. Choice is a dummy variable equal to one if a pupil chose (or brought in their packed lunch) a fruit or vegetable portion on a given day. Try is a dummy variable equal to one if the pupil eats at least some of a fruit or vegetable portion on that day. "Eat more than half" is a dummy variable equal to one if the pupil eats at least some of a fruit or vegetable portion on that day. 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						
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
	[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	0.024	0.010	0.037	-0.033	0.033	-0.066**	0.126*

Incentive (=1) * Week 2-5	(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)
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

*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. Choice is a dummy variable equal to one if a pupil chose (or brought in their packed lunch) a fruit or vegetable portion on a given day. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6. 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 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						
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
	[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)
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

*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. Choice is a dummy variable equal to one if a pupil chose (or brought in their packed lunch) a fruit or vegetable portion on a given day. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6. 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 4: Effect on Trying for Overall Sample and Its Subgroups

	Dependent Variable (=1) if Student Tried a Healthy Item						
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5

	[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)
P-Value: Comp = Ind Incentive Week 2-5	0.251	***0.041***	0.730	***0.010***	0.418	***0.002***	0.247
P-Value:	0.286	***0.068***	0.807	***0.020***	0.464	***0.002***	0.378

Comp = Ind Incentive Week 2-5 (wild)							
P-Value: Comp = Ind Incentive Week 6	0.164	***0.054***	0.484	***0.012***	0.323	***0.000***	0.256
P-Value: Comp = Ind Incentive Week 6 (wild)	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

*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. Try is a dummy variable equal to one if the pupil eats at least some of a fruit or vegetable portion on that day. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6. 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 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						
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
	[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)
P-Value:	0.167	***0.067***	0.463	0.239	0.192	***0.000***	0.662

Comp = Ind Incentive Week 2-5							
P-Value: Comp = Ind Incentive Week 2-5 (wild)	0.188	***0.092***	0.527	0.484	0.206	***0.004***	0.743
P-Value: Comp = Ind Incentive Week 6	0.117	***0.047***	0.301	0.322	0.126	***0.000***	0.715
P-Value: Comp = Ind Incentive Week 6 (wild)	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

*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. Try is a dummy variable equal to one if the pupil eats at least some of a fruit or vegetable portion on that day. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6. 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 6: Effects on Choice and Try Over Treatment

Weeks on Sample with Week 1 less than 100% Choice

	Dependent Variable (=1) if Student			
	Chose or Tried a Healthy Item			
	Choice	Choice	Try	Try
	[1]	[2]	[3]	[4]
Competition (=1) * Week 2-5	0.043	0.112	0.120	0.182
	(0.085)	(0.093)	(0.110)	(0.111)
	[0.649]	[0.330]	[0.346]	[0.192]
Individual Incentive (=1) * Week 2-5	-0.044	0.176**	-0.044	0.185**
	(0.200)	(0.064)	(0.201)	(0.073)
	[0.799]	[0.162]	[0.873]	[0.242]
Constant	0.327***	0.546***	0.223***	0.589***
	(0.050)	(0.038)	(0.045)	(0.042)

Days of the Week Used	Wed-Fri	Wed-Fri	Wed-Fri	Wed-Fri
Sample Used	Missed	Not Missed	Missed	Not Missed
P-Value: Comp = Ind Incentive	0.664	0.557	0.435	0.984
P-Value: Comp = Ind Incentive (wild)	0.677	0.661	0.490	1.007
Observations	876	1,982	887	1,924
R-squared	0.029	0.080	0.035	0.081
Number of pupils	158	202	157	203

*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. Choice is a dummy variable equal to one if a pupil chose (or brought in their packed lunch) a fruit or vegetable portion on a given day. Try is a dummy variable equal to one if the pupil eats at least some of a fruit or vegetable portion on that day. The incentives were in place during weeks 2 to 5 (inclusive). 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 columns [1] and [3] 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 [2] and [4] includes only those children who had eaten at least one fruit or vegetable on Monday or Tuesday during the given week.*

Table 7: Long Run Effect on Choice for Overall Sample and Its Subgroups

	Dependent Variable (=1) if Student Tried a Healthy Item						
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5

	[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
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

*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. Choice is a dummy variable equal to one if a pupil chose (or brought in their packed lunch) a fruit or vegetable portion on a given day. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6 or 6 months later.*

Table 8: Long Run Effect on Try for Overall Sample and Its Subgroups

	Dependent Variable (=1) if Student Tried a Healthy Item						
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
	[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	0.875	0.911	0.681	0.168	0.759	0.304	***0.010***

Months (wild)							
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]
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	0.523	0.513	0.711	0.817	0.581	0.468	0.913

(wild)							
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

*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. Try is a dummy variable equal to one if the pupil eats at least some of a fruit or vegetable portion on that day. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6 or 6 months later.*

Table 9: Food Knowledge

	Dependent Variable: Change in Food knowledge Test Score						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
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	0.827	0.852	0.798	0.754	0.843	0.806	0.853

Score							
Mean of Dependent Variable	0.022	0.008	0.038	0.061	0.013	0.024	0.020
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

*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 10: Food Knowledge on Sample with Week 1 less than 100% Choice

	Dependent Variable: Change in Food knowledge Test Score						
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition	-0.011	-0.032	0.017	-0.133	-0.003	-0.113	0.061***

(=1)	(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
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

*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

Figure A2: Example food knowledge test

Appendix B: Additional Figures and Tables

Not for Publication

Table 2A: Tests for Differences Between Subgroups

	Column	Column	Column
	[2] = [3]	[4] = [5]	[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	0.571	0.316	***0.076***

(=1) * Week 2-5 (wild-p)			
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

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. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6.

Table 3A: Tests for Differences Between Subgroups

	Column	Column	Column
	[2] = [3]	[4] = [5]	[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) *	0.473	0.951	0.570

Week 6			
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

The table contains the p-values for the tests of whether the coefficient on the variables in Table 3 for the two columns listed are equal. Wild p-values shown are estimated following Cameron, Gelbrach, Miller (2008) using 1000 replications. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6.

Table 4A: Tests for Differences Between Subgroups

	Column	Column	Column

	[2] = [3]	[4] = [5]	[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	Male	Non-FSM	Year 5

Column Heading			

The table contains the p-values for the tests of whether the coefficient on the variables in Table 4 for the two columns listed are equal. Wild p-values shown are estimated following Cameron, Gelbrach, Miller (2008) using 1000 replications. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6.

Table 5A: Tests for Differences Between Subgroups

	Column	Column	Column
	[2] = [3]	[4] = [5]	[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

The table contains the p-values for the tests of whether the coefficient on the variables in Table 5 for the two columns listed are equal. Wild p-values shown are estimated following Cameron, Gelbrach, Miller (2008) using 1000 replications. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6.

Table 7A: Tests for Differences Between Subgroups

	Column	Column	Column
	[2] = [3]	[4] = [5]	[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

The table contains the p-values for the tests of whether the coefficient on the variables in Table 7 for the two columns listed are equal. Wild p-values shown are estimated following Cameron, Gelbrach, Miller (2008). The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6 or 6 months later.

Table 8A: Tests for Differences Between Subgroups

	Column	Column	Column
	[2] = [3]	[4] = [5]	[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	Male	Non-FSM	Year 5

Column Heading			

The table contains the p-values for the tests of whether the coefficent on the variables in Table 8 for the two columns listed are equal. Wild p-values shown are estimated following Cameron, Gelbrach, Miller (2008) using 1000 replications. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6 or 6 months later.

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: Descriptive Statistics of Local Education Authorities

	(1) Participate	(2) Collaborate but not participate	(3) Did not collaborate	(4) p-value
Contacted on Friday	0.75	0.41	0.47	0.163
Contacted by J James	0.58	0.53	0.50	0.795
Household Income/100	6.10	7.23	6.72	0.138
% FSM	0.15	0.21	0.17	0.164
Number of Schools/100	2.27	1.50	1.35	0.037
% 5 Fruit & Veg a day	24.5	25.9	26.0	0.603

% Overweight & Obese reception	23.6	23.2	23.0	0.714
% Obese reception	9.94	10.4	9.74	0.330
% Overweight & Obese yr6	33.1	34.4	33.0	0.309
% Obese yr6	18.5	20.1	18.7	0.180
Smoking	25.6	24.7	24.6	0.794
Binge Drinking	20.2	17.1	18.1	0.195
Key stage 1: Avg point score	0.15	0.15	0.15	0.879
Key stage 2: Avg point score	0.28	0.28	0.28	0.894
Per pupil spending 2010/11	4307	4806	4486	0.109
% change in per pupil spending 2010/11	0.04	0.04	0.04	0.778
% LA spending change 2010/11	-0.05	-0.05	-0.05	0.689
Female CEO of	0.25	0.31	0.25	0.859

the council				
Female Director of Children Services	0.67	0.41	0.51	0.405
Female Leader of Healthy Schools	0.75	0.88	0.82	0.659
% of Labour Councillors	0.32	0.39	0.33	0.650
% of Conservative Councillors	0.43	0.43	0.42	0.983
Labour controlled council	0.25	0.29	0.29	0.958
Conservative controlled council	0.42	0.47	0.41	0.878
Ofsted Score	2.34	2.31	2.28	0.405
Ofsted Health Score	1.69	1.72	1.70	0.707
Catering pp/100	0.80	0.83	0.82	0.989
Energy costs	0.69	0.68	0.65	0.610

pp/100				
Total school Income pp/1000	4.30	4.52	4.32	0.674
Teaching costs pp/1000	2.10	2.18	2.13	0.756

P-values in column 4 come from a test of equality of the 3 group means between local authorities that participated, those that collaborated (by providing names of schools), and those that did not collaborate. Local authorities were randomly contacted on two days on Friday 2nd July and Monday 5th July and by J. James or M. Belot. Income is the average weekly total household income (£) divided by 100, FSM is the percentage of children who are eligible for free school meals. % Eat 5 Fruit & Veg a day is the proportion of adults defined to be consumers of 5 or more fruit and vegetables if they had reported that they had consumed 5 or more portions of fruit and vegetables on the previous day. Binge drinking is the proportion of adult binge drinkers defined if they reported that in the last week they had drunk 8 or more units of alcohol if they were a man, or 6 or more units of alcohol if they were a woman, on any one day or more. Smoking is the proportion of individuals in a local authority who reported that they were a 'current cigarette smoker' in the Health Survey for England. Overweight and Obese reception is the percentage of pupils in the local authority who were overweight or obese when they entered primary school aged 4 or 5. Year 6 is the final year of school when the pupils are aged 10 or 11. The average point score (APS) of the key stage 1 test and key stage 2 point score are for tests taken in primary school. The points are awarded per subject per pupil along the following lines: working below the level of the test or not awarded 15, level 2 receives 15 points, level 3 gets 21, level 4 gets 27 and 33 points is allocated for level 5. The APS is then calculated using the following: (Total points for English + Total points for maths + Total points for science) / (Total number of eligible pupils for each subject). This is then rescaled by dividing by 100. Per pupil spending in 2010/11, the yearly increase in per pupil spending, and the overall change in the spending of the local authority. Labour Party and Conservative Party councillors on the council defined at the most recent election since July 2010. Ofsted is (the government school inspector) average score of the schools in the local authority. Schools are inspected and judged on the following question: "How effective, efficient and inclusive is the provision of education, integrated care and any extended services in meeting the needs of learners?" With ratings given of: 1. Outstanding 2. Good 3. Satisfactory 4. Inadequate. Ofsted Health Score is based on the following question: "Learners are encouraged and enabled to eat and drink healthily" using the same 1 to 4 scale. Average catering costs (including staff costs), energy, teaching and total school income are per pupil averages at the local authority level and are rescaled as indicated.

Table B2 Comparison of participating schools from the pool of selected schools

	Experiment	Dropped	p-value of difference
% Girls	0.48	0.49	0.802
Number of pupils	207	279	0.322
% FSM Eligible	15.9	15.8	0.849
% FSM Take	13.9	13.7	0.802
Total School Income per pupil/1000	4.17	4.16	0.641
Absenteeism (% on census day)	0.05	0.05	0.682
Catering costs per pupil/100	0.96	0.73	0.303
% English and Maths above level 4 KS2	0.76	0.76	0.949
Average point score Maths and English	0.28	0.28	0.396
Food for life	0.31	0.21	0.501
Ofsted Score	2.09	2.29	0.521
Ofsted health Score	1.53	1.43	0.604

FSM Medium	0.29	0.29	0.975
FSM Low	0.65	0.64	0.988
Teaching Costs per pupil/1000	2.05	2.17	0.246
Energy costs per pupil/100	0.64	0.87	0.961
Competition treatment	0.29	0.43	0.368
Individual treatment	0.32	0.36	0.822
Control	0.39	0.21	0.260
Schools	31	15	

Columns 1 and 2 show the mean values at the school level. Column 3 is the p-value of (Prob>z, where z is the test statistic) from an Mann-Whitney U test. Ofsted is (the government school inspector) average score of the schools in the local authority. Schools are inspected and judged on the following question: "How effective, efficient and inclusive is the provision of education, integrated care and any extended services in meeting the needs of learners?" With ratings given of: 1. Outstanding 2. Good 3. Satisfactory 4. Inadequate. Ofsted Health Score is based on the following question: "Learners are encouraged and enabled to eat and drink healthily" using the same 1 to 4 scale. Average catering costs (including staff costs), energy, teaching and total school income are per pupil averages at the local authority level and are rescaled as indicated. FSM Band - The three broad bands used to group pupils eligible for FSM are: Low: less than 20%, Medium: 20.01-35% and High: greater than 35% (omitted category). Columns (1)-(3) present estimates using whether a school was selected by the LEA. Column (3) excludes "Avg Eng/Math Score" but uses the same sample in column (2). Column (4) and (5) use whether a school started and completed the experimental intervention.

Table B3: Effect on Eating More than Half for Overall Sample and Its Subgroups

	Dependent Variable (=1) if Student Ate More than Half a Healthy Item

	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
	[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)
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

*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). "Eat more than half" is a dummy variable equal to one if the pupil eats at least half of a portion of fruits or vegetables on that day. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6. 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 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						
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
	[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)
P-value for Competition=Individual	0.199	0.104	0.420	0.391	0.183	***0.001***	0.883

for Wks 2-5							
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

*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). "Eat more than half" is a dummy variable equal to one if the pupil eats at least half of a portion of fruits or vegetables on that day. The incentives were in place during weeks 2 to 5 (inclusive) but were not in place during week 6. 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 Attendance On Overall Sample and Its Subgroups

	Dependent Variable (=1) if Student Attended School						
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5

	[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)
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

*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 Attendance for Sample with Week 1 less than 100% Choice and Its Subgroups

	Dependent Variable (=1) if Student Attended School						
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
	[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)
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

*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 B7: Effect on Bringing Packed Lunch On Overall Sample and Its Subgroups

	Dependent Variable (=1) if Student Brought a Packed Lunch						
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1) *	0.000	0.000	-0.001	0.014	0.001	0.008	-0.014

Week 2-5	(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)
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

*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 B8: 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						
	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1) *	0.007	-0.021	0.040	0.033	-0.000	0.020	-0.019
Week 2-5	(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) *	-0.004	-0.076	0.080**	-0.006	0.003	-0.039	0.005
Week 6	(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	0.036	0.054*	0.022	0.007	0.053*	0.060	0.027
(=1) * Week 2-5	(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	0.018	0.044	-0.003	-0.017	0.048	-0.039	0.050
(=1) * Week 6	(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)
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

*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.*