HABIT FORMATION IN CHILDREN:

EVIDENCE FROM INCENTIVES FOR HEALTHY EATING

George Loewenstein
Joseph Price
Kevin Volpp

We present findings from a field experiment conducted at 40 elementary schools involving 8,000 children and 400,000 child-day observations, which tested whether providing short-run incentives can create habit formation in children. Over a three or five week period, students received an incentive for eating a serving of fruits or vegetables during lunch. Relative to an average baseline rate of 39%, providing small incentives doubled the fraction of children eating at least one serving of fruits or vegetables. Two months after the end of the intervention, the consumption rate at schools remained 21% above baseline for the three-week treatment and 44% above baseline for the five week treatment, a significant difference. These findings indicate that short-run incentives can produce changes in behavior that persist after incentives are removed and support the natural intuition that longer interventions produce more persistent habits.

JEL Codes: J13, I18, I28

---

* Joseph Price is the corresponding author: 162 FOB; Brigham Young University; Provo, UT 84602; 801-422-5296; joseph_price@byu.edu. The authors are grateful to the USDA Economic Research Service for research funding and for the many BYU undergraduates and PTA parents that helped as research assistants on the project. This paper has benefitted from the helpful comments by Sandy Black, Dan Hamermesh, Patrick Nolen, Devin Pope, Bob Slonim, Coady Wing, Jingsan Zhu and seminar participants at the Cornell University, Harris School, University of Illinois at Chicago, University of Sydney, UT-Austin, Iowa State, University of Utah, University of Oregon, University of Iowa, and Brigham Young University.

Currently, there is vigorous debate about when it is either effective or appropriate to incentivize positive behaviors in children. Opponents of the use of incentives argue that extrinsic rewards crowd out intrinsic motivation and results in outcomes being worse after the end of the incentive period than prior to the introduction of rewards (Deci, Koestner, and Ryan 1999), and there is, indeed, evidence of such effects in studies conducted by economists (see Frey and Jegen 2001 for a review). However, arguments against the use of incentives overlook the role that habit formation can play in promoting long run behavioral change. Dictionary.com defines a habit as "an acquired behavior pattern regularly followed until it has become almost involuntary." If this habit formation process occurs during and as a result of the provision of incentives for engaging in a behavior, then short-term efforts that encourage children to engage in a particular activity can, if sufficient to overcome any crowding out of intrinsic motivation, result in positive behavior change even after the incentives are removed.

In this paper, we examine the role of incentives in promoting healthy eating behaviors in children, in particular consumption of fruits and vegetables. Inadequate consumption of fruits and vegetables is widely seen as an important contributor to suboptimal health worldwide and increases the risk for cardiovascular diseases, stomach cancer and colorectal cancer. However, achieving high rates of fruit and vegetable consumption among children has proved a considerable challenge and has been the focus of a number of recent school-based interventions.

We implemented an incentive program at 40 elementary schools in Utah in which children could receive a special token each day as a reward for consuming at least one serving of fruits or vegetables. The tokens were worth $.25 and could be spent at the school store, school carnival, or book fair. Schools were randomly assigned to implement the incentives for a period of either 3 or 5 weeks. We observe detailed fruit and vegetable consumption data at these schools before, during, and for two months after the intervention ends. This experimental design allows us to examine whether the increase in fruit and vegetable consumption that we observe during the incentive period persists once the incentives are removed. The central, and novel, focus of the study is whether a longer intervention results in larger post-incentive behavioral change – i.e. greater habit-formation.

BACKGROUND

The results of this paper complement a few other recent studies that examine the impact of incentives on children's in-school food choices. Just and Price (2013) provided incentives for five days and found lingering effects during the first two weeks after the intervention, but these did not persist four weeks after the intervention. Belot, James and Nolen (2013) provided students with stickers for choosing healthy items over a 4 week period. In addition, they provided a bonus gift under various conditions. In the piece-rate condition, each child received a bonus gift if they chose at least 4 healthy items in a week and in the competition condition, the child(ren) who collected the largest number of stickers received the bonus gift. Although the incentives only rewarded choice rather than consumption, there was noticeable change in consumption both one week and six months after the end of the rewards period. List and Samek (2013) provided low income school students with a small prize as a reward for choosing a healthier snack (dried fruit) over a less healthy snack (a cookie). They varied whether the incentives were framed either as gains or losses and whether or not the incentive was accompanied by a health message. They observed a large impact of incentives on the children's choices and they observed habit persistence when incentives were removed which was greater with the gain frame and when incentives and messaging were combined.

Studies of habit formation in domains other than school children's food choices have yielded mixed effects. Charness and Gneezy (2009) randomly assigned college students to one of three conditions: no incentive for gym attendance, $25 to attend the gym one time, or $25 to attend the gym one time plus $100 to attend the gym another 8 times. Their key finding was that, consistent with habit formation, subjects in the high incentive treatment group had higher gym attendance (about .6 more visits per week) during the post-incentive period than those in the low incentive and no incentive groups.

In a replication and extension of this study, Acland and Levy (2013) observed a smaller post-incentive effect (.26 visits per week), and found that the effect decayed over the course of the winter vacation and was highly concentrated in the upper tail of the post-treatment attendance distribution. Royer, Stehr and Sydnor (2012) also tested a similar intervention using adult workers at a Fortune 500 company and additionally tested the impact of giving workers access to a self-funded commitment contract. They found a weak persistence of gym use after the incentive was withdrawn among those provided with an incentive alone (16% of the increase in attendance during the incentive period), but substantially greater persistence (47%) among those who were provided access to the commitment contract.

In some contexts, persistence of behavior change may be easier to achieve than others. Volpp et al. (2009) randomized smokers into a treatment group which offered a $750 incentive ($100 for completion of a program, $250 for short-term cessation, $400 for long-term cessation). This incentive resulted in a quit rate of 14.7 percent in the intervention group compared to 5.0 percent in the control group at 12 months. Six months after the long-term incentives were discontinued, the quit ratio was 2.6 (9.4 percent versus 3.6 percent), suggesting that if incentives are effective in helping an individual to stay smoke-free for 12 months, there is a reasonable chance they will develop habits that increase their likelihood of remaining smoke-free when incentives are withdrawn. In contrast, weight loss interventions have typically shown less evidence of habit formation. In two studies testing the use of lottery incentives and deposit contracts for weight loss (Volpp et al., 2008; John et al., 2011), incentives were highly effective in motivating weight loss during the incentive period, but participants regained most of the weight they had lost once the incentives ended.

METHODS

We conducted a field experiment at 40 elementary schools in Utah involving 8,000 students in grades 1-6. The data was collected over a period of 18 months, from January 2012 to June 2013, with seven schools participating during the winter 2012 semester, 10 schools the following fall, and 25 schools in 2013.[1] The schools in our study provide students at lunchtime with a choice of a main entrée from two or three options, and students are allowed to choose as many additional items as they want from a selection of fruits, vegetables, and other side dishes. It was important that we could accurately measure the number of servings each student placed on their tray and the number they consumed, so most of the fruits and vegetables came in special cups while others, such as bananas or oranges, were quantified by the leftover peel or core.

Research assistants stood by the trash cans in each cafeteria and recorded the number of fruits and vegetables both taken and consumed by each child by observing each child's tray as they exited the lunch room. In cafeterias where there were multiple trash locations, at least one assistant was stationed at each location. For cases where trays were empty but students were seen carrying out fruit or vegetable items for later consumption, students were marked as not having eaten that item. Prior to the rewards period, baseline data were collected at each school for two

[1] The new lunch guidelines were put in place at the start of the 2012-2013 school year so all of the schools in our sample participated in the experiment either completely before or completely after the change in guidelines (particularly the requirement that every child take a serving of fruits or vegetables). All of our analysis includes school fixed effects to account for any differences across schools in how the guidelines were implemented or other differences in their lunch program.

weeks. To avoid influencing the children's behavior in the baseline observations, our data collectors did not mention fruits or vegetables when asked what they were doing, but, if asked, rather stated that they were simply collecting data about school lunches.

During the rewards period, children receiving a school lunch who ate at least one serving of fruits or vegetables received a special coin which had a picture of an apple and a carrot on it. These coins had a value to the students of twenty-five cents and could be redeemed at a school store, book fair, or school carnival. We used redeemable tokens instead of cash in response to a concern expressed by some school principals that children might use the money to purchase candy or other junk food after school. At the end of the program, each school received a check for the value of tokens that were redeemed, providing the double benefit to schools of encouraging healthy eating and providing additional funds for the school and PTA.

At the start of the rewards period, an announcement about the program was included in a newsletter sent home to parents, and information about the program was provided in the school's morning announcement (just prior to the start of the rewards period). Reinforcing these announcements, the research assistants handing out the tokens were instructed to explain to students why they were distributing the tokens and also reminded children who had not eaten a full serving of fruits or vegetables that if they went back and finished their fruit or vegetable they could receive a token. Thus the change behavior during the incentive period may result from both the direct effect of the incentives as well as any effects operating through the presence and interaction with the data collectors.

During the token period, some schools expressed concerns that some children might be cheating by hiding their vegetables in their milk cartons or throwing food on the floor in order to receive a token. Schools provided an announcement to students about the importance of honesty in the token program and warned that cheating would result in the end of the program for the school. In addition, since our data collectors were in the cafeteria during the entire lunch period, they were able to check for food being thrown on the floor or to ask the students personally if they did indeed eat an item before handing them a token. It is important to note, however, that, while any cheating will bias upwards the estimates of consumption during the incentive period, it should have no effect on our estimates of habit formation, since those estimates are based on the measures recorded after the end of the token period.

During most of the week, the tokens were only available to students who purchased or received a school lunch. This restriction was based on the fact that the grant funding our research was focused on making improvements to the school lunch program and because our data collection approach is not well suited to measure the fruit and vegetable consumption of students with a sack lunch. As an accommodation to these children, we made the tokens available to all students who consumed a fruit or vegetable on Fridays but did not include children with a sack lunch in any of our data collection. It is possible that students may have switched from getting a sack lunch to getting a school lunch on token days. It is likely that it would be the students most likely to eat fruits and vegetables anyways that would be the switch in response to the incentives.[2] These switchers may bias upwards our estimates of the effects during the token period but similar to the issue of cheating, they should not have any effect on our estimate of habit formation.

The 40 schools that participated in our experiment expressed willingness to implement an incentive program at their school for up to five weeks. Schools that elected to participate were randomly assigned to have the rewards in place for either three weeks or five weeks. To ensure

[2] Just and Price (2013) find that their rewards program increased the fraction of children eating school lunch that day by 4%. They note that if all of the children who switched over on the incentive days were already eating a serving of fruits or vegetables with their sack lunch that the effect of the incentive would need to be scaled down by one eighth.

that we would have similar schools in each of our treatment groups, we stratified the randomization on two school characteristics: the baseline consumption rate at the school and the fraction of children at the school who are on free and reduced price lunch (FRPL rate).

Our primary outcome measure is an indicator for whether or not the child ate at least one serving of fruits or vegetables. This was the criterion we used to determine whether a student received a reward or not. For each student we recorded the number of servings that they took and how many they actually ate. Measuring actual consumption was key since the majority of our schools participated in the field experiment after the implementation of the new lunch guidelines that require that students place at least one serving of fruits or vegetables on their tray. Our rewards program was designed to counteract the fact that these new guidelines were leading to such a large number of fruit and vegetables being thrown away. We also report results for a secondary measure: the number of servings of fruits and/or vegetables actually consumed per student.

Our final sample includes 40 elementary schools, 22 of which had the rewards period in place for 3 weeks and 18 which had it in place for 5 weeks. For our analysis, we use the student-day as the unit of observation but cluster all of our standard errors at the school level. Our combined sample has 403,922 child-day observations including the baseline, incentive, and follow-up periods. Running our analysis at the child-level allows us to include controls for student characteristics such as gender or grade. Our data collection procedure and IRB restrictions made it impossible to collect information on each student's identity so we are unable to include any student fixed effects. We do include both school fixed effects and day of the week fixed effects in our analyses.

RESULTS

In Table 1, we use data from the Common Core of Data to provide some basic characteristics of the schools in our sample. We also provide the p-value of the t-test for whether the characteristics differ between the two treatment groups. At the bottom of Table 1, we also provide some measures from our baseline data collection. Among the students at the schools in our sample, about 39% were eating at least one serving of fruits or vegetables and the average student was eating about 0.59 servings of fruits or vegetables every day. We find that none of the differences between the two treatment groups at baseline are statistically significant, suggesting that the block randomization approach was successful in balancing the characteristics across the two treatment groups.

Figure 1 provides the fraction of children eating at least one serving of fruits or vegetables during each of the four periods of the study separately for the two treatment groups. The results in this figure provide three general patterns. First, the incentives produced a very large change in the fraction of children eating fruits or vegetables during lunch (almost doubling the fraction eating at least one serving). Second, the high rates of fruit or vegetable consumption decreased after the end of the incentive period but remained at a level significantly higher than the baseline period. Third, the post-incentive fruit and vegetable consumption patterns were higher for schools assigned to the five-week treatment group than those assigned to the three-week treatment group.

Figure 2 provides a more disaggregated look at the consumption rates during the baseline and incentive period. Since data was collected at each school for ten days during the baseline period, we provide the average across all schools for these ten days. We then provide the average consumption across the 15 days at the three-week school and the 25 days at the five-week schools. This graph shows that consumption rates during the baseline period were relatively stable followed by a large and immediate change once incentives were in place. There is also an incremental increase in consumption during the first few days of the incentive program suggesting either an increase in awareness or some students changing their mind about wanting to earn a reward (possibly through the influence of peers).

Table 2 provides a similar descriptive view of change in behavior in each of the treatment groups. We split the incentive period up into two periods (weeks 1-3 and weeks 4-5) to compare consumption rates during the same window of time for the two treatment groups. We provide a separate row for the last week of the incentive period (this row is not mutually exclusive of the other two rows). We also present the results for an alternative outcome measure, the number of servings of fruits and vegetables consumed per student. The results in this table provide the same insights as Figure 1 but also show that slightly higher consumption rates during the incentive period at the five-week schools occurred even during the first three weeks (though the small difference during the incentive period is not statistically significant). The results in this table also show that the change in behavior was very similar over the course of the incentive period indicating that the effects of incentives did not fade out as they were left in place longer.

Both interventions significantly increased the consumption of fruits and vegetables (3-week intervention: 39.9% at baseline, average of 76.4% during intervention, p-value for difference <0.01; 5-week intervention: 37.6% at baseline, average of 79.5% during intervention p-value for difference <0.01). After the incentive period ended, the fraction of children eating at least one serving of fruits or vegetables decreased, but remained at a level about 10 percentage points above the baseline level in the 3-week intervention and 16.4 percentage points above baseline in the 5-week treatment (representing increases of 25.1% and 43.6%, respectively). These results indicate that the intervention did produce a meaningful change in post-incentive behavior for both treatment groups. This result is encouraging because it suggests that this type of incentive program did not have the type of ‘crowding out of intrinsic motivation’ effect that has been periodically raised as a concern about the use of incentives, or that such an effect, if present, was dominated by a stronger habit formation effect.

In Table 3 we provide the regression-based analog of the results in Table 2 that control for child, gender and grade and include day of week and school fixed effects. The regression-based estimates are very similar to the raw differences observed in Table 2, which is expected given the random assignment into treatment groups. For both incentive groups, we find that the increase in consumption was large and statistically significant both during the incentive period and up to two months after the end of the incentives.

Accounting for the persistent behavior during the two months after the incentive period also dramatically improves the cost effectiveness of the incentive program. Focusing on just the incentive period indicates that the intervention cost about 50 cents for each additional child induced to eat a serving of fruits or vegetables (52.1 at the 3-week schools and 47.9 at the 5-week schools). Once we include the additional consumption that occurs after the incentive period the cost per additional child eating a serving of fruits or vegetables drops to about 28 cents (29.0 cents at the 3-week schools and 28.4 at the 5-week schools). These estimates would have undoubtedly decreased even more if we had continued to measure consumption data even longer than two months after the end of the incentive period.

Another important question is whether the longer incentive period led to higher post-intervention fruit and vegetable consumption, indicative of greater habit formation. Table 4 presents regressions that pool the two treatment groups together and estimate the interaction term between the length of the incentive and each of the periods of the study (e.g. incentive, 1-month follow-up, and 2-month follow-up). The coefficients on the interaction terms provide evidence that the 5-week intervention produced incrementally greater persistence (about a tenth of a serving per student per day) in the month following the intervention. At 2 months post-intervention, there is some suggestive evidence of greater persistence in the 5-week intervention group (54.0% vs. 48.1%) though the p-value on this difference is 0.262.

Our study presents a common situation in randomized field experiments in which there is a very large sample but a much smaller number of randomization units. Various approaches have been developed to estimate appropriate standard errors that take into account the intra-class correlation between observations from the same randomizing unit. The standard errors reported in Table 4 have all been clustered at the school level. We also implemented a set of alternative approaches for calculating the standard errors used in past studies. These approaches include the cluster generalization of the wild bootstrap described by Cameron, Gelbach, and Miller (2008), the paired bootstrap method used by Prescott and Rockoff (2008), and the method of Generalized Estimating Equations (GEE) developed by Liang and Zeger (1986).

All of these alternative approaches (reported in Appendix Table 1) provide very similar standard errors as the ones we report in Table 4. The standard errors using the GEE approach tend to produce the most precise estimates, and under this approach the difference between the three and five week schools one month after the end of the incentives would be statistically significant at the 5% level. If we did not cluster the standard errors at the school level, we would have standard errors that are about 10 times smaller suggesting that future researchers might consider ways in which they can increase the number of clusters even if it requires reducing the overall sample. Some options might include having just one or two grades per school participate or possibly restricting the sample to smaller schools.

A final approach to statistical inference in this situation is to use permutation inference. Under the null hypothesis that there is no difference in follow-up period between 3-week and 5-week schools, mislabeling schools as a three or five-week school would have no effect on the estimated coefficient of the interaction terms in Table 4. We randomly generated 10,000 permutations of the labeling across the 40 schools (holding constant the number of schools assigned the label of five week schools) and estimated the same model as Table 4. We find that only 246 of these permutations had a coefficient on the interaction between the five week treatment and one-month follow-up period that was larger than the coefficient reported in Table 4. This provides a p-value for the two-sided test between 5-week and 3- week treatments of 0.049. In addition, the permutations that had the highest estimated coefficients were those in which the highest fraction of schools were assigned the correct label providing additional evidence that the 5-week treatment actually had a larger effect after the end of the incentive period.[3]

DISCUSSION

The results of this paper are based on a large field experiment at 40 elementary schools in which children received a small incentive for consuming fruits and vegetables as part of their school-provided lunches. We find that these small incentives produced a dramatic increase in fruit and vegetable consumption during the incentive period and that this change in behavior was sustained for at least two months after the incentives stopped. We also find suggestive evidence that a longer intervention period produced a more sustained response once the rewards were removed.

One question raised by this and other related studies is the mechanism that led to behavior persistence once incentives were removed. At least three mechanisms are possible.

---

[3] Heckman et al (2010) use a permutation-based inference approach in their examination of the Perry Preschool Program and highlight some of the advantages of this approach.

One, which could be considered ‘classic’ habit formation, suggests that students became used to eating fruits and vegetables during lunch and this became an automatic pattern of behavior. A second mechanism is that consuming the fruits and vegetables may have led to either a discovery of pre-existing tastes or a change in tastes, the latter consistent with prior research which shows that repeated exposure to specific items can influence an individual’s food preferences (Birch & Marlin 1982). The third mechanism is that making fruit and vegetable consumption more ‘popular’ (albeit with the help of an incentive) may have shifted social norms around fruit and vegetable consumption such that kids would be less likely to cast aspersions on other kids who ate fruit and vegetables at lunch.

An observation related to why habit formation seems more likely in some settings than others in previous studies (smoking, food choice as opposed to losing weight) is that there are some settings where new routines develop and where the environmental setting reinforces the behavior change choice (smoking or eating healthy food in a cafeteria) in contrast to settings where environment is generally stacked against the behavior change (weight loss). A daily routine around a specific task such as getting a tray each day at the same time and changing one component of what is on the tray is also far simpler than trying to change a whole host of elements required to lose weight.

One other mechanism that could explain behavior persistence in this context was the presence of data collectors in the lunchroom after the incentive was removed. Since they were the ones who distributed tokens to the students and encouraged them to eat their fruits and vegetables for several weeks, perceived social pressure on the part of the student may be enough to induce them to continue to eat fruits and vegetables. An important piece of evidence that points against this mechanism are the results of Just and Price (2012) who implemented a program very similar to ours but left the rewards in place for only five days. They found that following the end of the incentive period consumption rates went back to their baseline levels, suggesting that the presence of the data collectors alone isn't sufficient to create the false appearance of habit formation.

None of the treatments showed any evidence of a rebound effect, where consumption patterns after the incentive period dropped to levels lower than pre-treatment levels. This has been a concern raised by some psychologists about the use of extrinsic motivators to change behavior in children (Deci, Koestner, & Ryan 1999), although concerns have been raised about the robustness of these results as well as the vulnerability to alternative interpretation (Cameron & Pierce 1994). In the current experiment, we cannot rule out the possibility that such effects occurred but the fact that we observed persistence instead of rebound, suggests, at a minimum, that any such effects were exceeded by the influence of habit formation.

Results from our study reinforce those from earlier research showing that the use of small incentives is an effective way of encouraging children to eat more fruits and vegetables and that these induced changes in behavior persist after the incentives are no longer being offered (Belot et al 2013; List and Samek 2013). We also find suggestive evidence that longer intervention periods lead to greater persistence of behavior change. While the habit formation process that we observe in this study may be most germane to food choices among children, there are many other positive health behaviors for which sustaining a period of active involvement can result in the behavioral change persisting even after the incentive is removed and where an approach similar to the intervention described here could be effective.

REFERENCES

Acland D. and Levy M. 2013. Naiveté, Projection Bias, and Habit Formation in Gym Attendance Working paper series, Social Science Research Network.

Belot, M., James J. and Nolen P. 2013. Changing Eating Habits- A Field Experiment in Primary Schools. Working paper, University of Edinburgh

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. *Appetite*, 3(4): 353-360.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. 2008. Bootstrap-Based Improvements for Inference with Clustered Errors *The Review of Economics and Statistics*, *90*(3): 414-427.

Cameron, J. and Pierce, W.D., 1994. Reinforcement, Reward, and Intrinsic Motivation: A Meta-Analysis. *Review of Educational Research* 64(3): 363-423.

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

Deci E.L., Koestner R., and Ryan R.M. 1999. A Meta-Analytic Review of Experiments Examining the Effect of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin*, 125(6): 627-668.

Heckman, J., Moon S., Pinto R., Savelyev P., and Yavitz A. 2010. Analyzing Social Experiments as Implemented: A Reexamination of the Evidence from the HighScope Perry Preschool Program. *Quantitative Economics,* 1(1): 1-46.

Just D. and Price J. 2013. Using Incentives to Encourage Healthy Eating in Children. *Journal of Human Resources,* 48(4): 885-872.

John L., Loewenstein G., Troxel A., Norton L., Fassbender J., and Volpp K. 2011. Financial Incentives for Extended Weight Loss: A Randomized, Controlled Trial. *Journal of General Internal Medicine*, 26(6): 621-626.

Liang, K. Y., & Zeger, S. L. 1986. Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika*, *73*(1): 13-22.

List, J.A. & Samek, S. 2013. The Behavioralist as Nutritionist: Leveraging Behavioral Economics to Improve Child Food Choice and Consumption. *National Bureau of Economic Research Working Paper 20132.*

Prescott, J., & Rockoff, J. 2011.Do Sex Offender Registration and Notification Laws Affect Criminal Behavior? *Journal of Law and Economics*, *53*: 161-206.

Royer, H., Stehr, M. F., & Sydnor, J. R. 2012. Incentives, Commitments and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company. *National Bureau of Economic Research Working Paper 18580*.

Volpp K., John L., Troxel A., Norton L., Fassbender J., and Lowenstein G. 2008. Financial Incentive–Based Approaches for Weight Loss: A Randomized Trial. *Journal of the American Medical Association*. 300: 2631-2637.

Volpp K.G., Troxel A.B., Pauly M.V., Glick H.A., Puig A., Asch D.A., Galvin R., Zhu J., Wan F., DeGuzman J., Corbett E., Weiner J., and Audrain-McGovern J. 2009. A Randomized, Controlled Trial of Financial Incentives for Smoking Cessation. *The New England Journal of Medicine*, 360(7): 699-709.

Table 1. Comparison of baseline characteristics across treatment groups.

	3 Weeks	5 Weeks	p-value
Number of students per school	663	629	0.55
	[190]	[150]	
Gender Ratio			
Male	50.6	50.6	0.97
	[2.6]	[1.8]	
Ethnicity			
White	79.4	75.4	0.47
	[14.7]	[20.3]	
Hispanic	14.5	18.5	0.41
	[13.1]	[17.5]	
Other	6.1	6.0	0.98
	[3.4]	[4.5]	
FRPL rate	40.5	45.3	0.43
	[16.2]	[22.2]	
Fraction of children eating at least one serving of fruits or vegetables	39.9	37.6	0.58
	[11.3]	[12.0]	
Amount of servings consumed per child	0.577	0.602	0.83
	[0.171]	[0.174]	
Number of schools	22	18	

Notes: Values listed above are averaged over all schools in each group. FRPL rate is the fraction of students who receive a free or reduced price lunch. Standard deviations are included in brackets. The p-value is based on a t-test for the difference in characteristics between the 3- and 5-week schools.

Table 2. Comparison of fraction of children eating a serving of fruits or vegetables by treatment group.

	Ate at least one serving		Amount of servings eaten	
	3 Weeks	5 Weeks	3 Weeks	5 Weeks
Baseline	39.9%	37.6%	0.577	0.602
Weeks 1-3 of intervention period	76.4%	79.9%	0.948	0.954
Weeks 4-5 of intervention period	-	79.0%	-	0.934
Last week of intervention period	75.4%	79.7%	0.906	0.927
1 Month After	49.6%	58.5%	0.658	0.762
2 Months After	48.1%	54.0%	0.648	0.716
N	191,719	212,203	191,719	212,203

Notes: The unit of observation is the student day. Both the 3- and 5-week interventions resulted in significant increases in fruit and vegetable consumption from baseline to the intervention period (p-values <0.01). None of the differences between the 3- and 5-week treatments at any of the specific time intervals are statistically significant at the 0.10 level. The p-value for the difference for "1 month after" is 0.109 in the first two columns and 0.110 in the last two columns.

Table 3. Impact of incentives on behavior after incentives are removed.

	Ate at least one serving		Number of servings eaten	
	3 Week	5 Week	3 Week	5 Week
Incentive	0.368**	0.411**	0.370**	0.348**
	(0.025)	(0.036)	(0.027)	(0.047)
1 month post-intervention	0.117**	0.205**	0.103**	0.171**
	(0.023)	(0.038)	(0.030)	(0.045)
2 months post-intervention	0.103**	0.148**	0.090**	0.122**
	(0.023)	(0.032)	(0.029)	(0.031)
Grade	0.005	0.001	0.005	0.001
	(0.003)	(0.003)	(0.003)	(0.004)
Male	-0.060**	-0.049**	-0.085**	-0.072**
	(0.006)	(0.009)	(0.008)	(0.011)
N	191,719	212,203	191,719	212,203

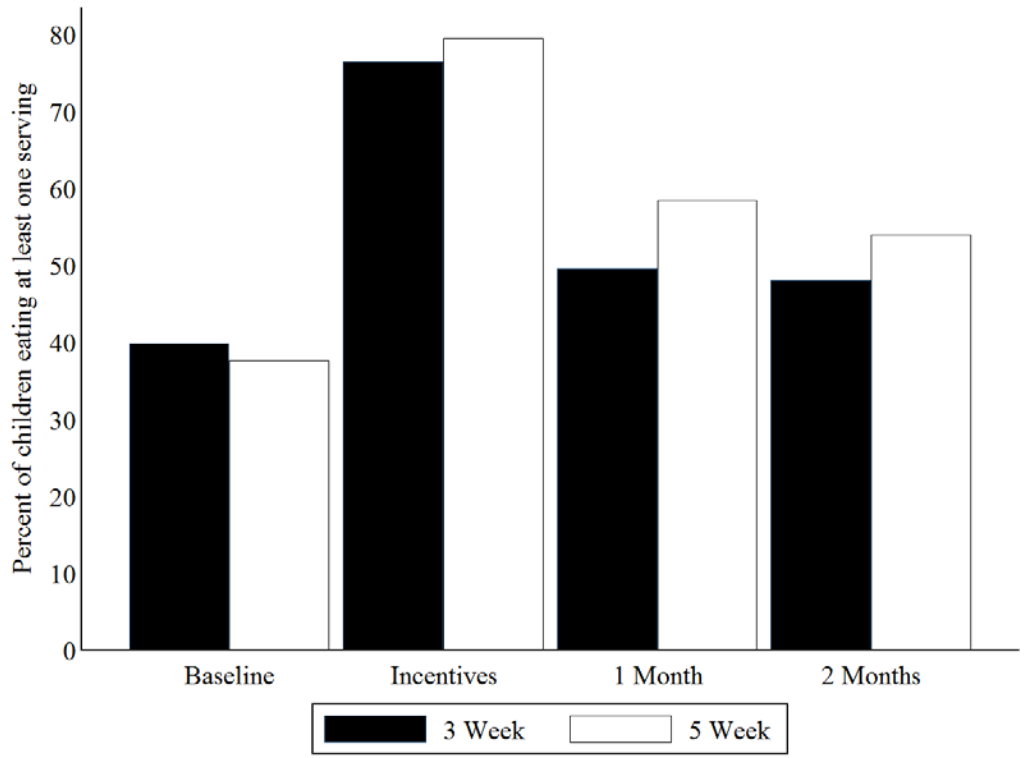
Notes: The unit of analysis is the student day. The regressions include school and day of week fixed effects and controls for the child's grade and gender. Standard errors are clustered at the school level. **, and * indicate statistical significance at the 1% and 5% levels, respectively.

Table 4. Impact of incentives on behavior after incentives are removed with interaction based on length of treatment period.

	Ate at least one serving	Number of servings eaten
Incentive	0.368**	0.369**
	(0.025)	(0.027)
1 month post-intervention	0.117**	0.104**
	(0.023)	(0.030)
2 months post-intervention	0.103**	0.090**
	(0.023)	(0.029)
Fiveweek*Incentive	0.042	-0.021
	(0.043)	(0.053)
Fiveweek*Month1	0.087†	0.067
	(0.044)	(0.053)
Fiveweek*Month2	0.044	0.031
	(0.039)	(0.042)
Observations	403,922	403,922

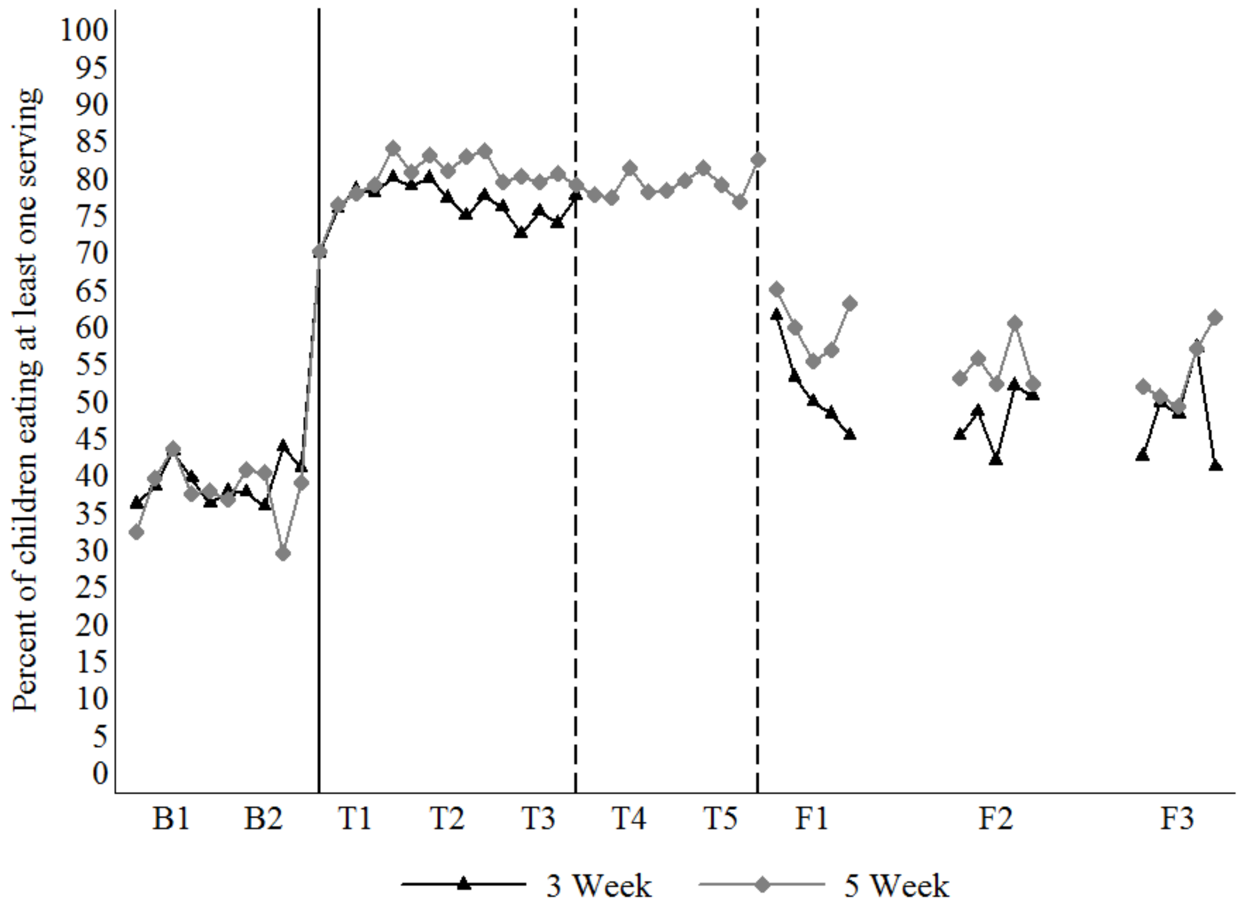
Notes: Each regression includes controls for gender, grade, day of week, and school fixed effects. Standard errors are clustered at the school level. **, *, and † signify statistical significance at the 1%, 5%, and 10% levels respectively.

Figure 1. Change in consumption before, during and after the incentives period.



Notes: The baseline period includes the ten days prior to the start of the incentives. "1 month" and "2 months" refer to the first and second month after the end of the incentive period.

Figure 2. Percentage of children eating at least one serving before and during the experiment.



Notes: This figure provides the fraction of children eating at least one serving of fruits or vegetables each day during the two weeks of baseline data collection and the 3-5 weeks rewards period. The solid vertical line indicates the start of the incentive period and the dashed vertical lines indicate the end of the 3- and 5-week incentive periods.

Appendix Table 1. Alternative approaches to estimating the standard errors in Table 4.

	Demeaned Regression	Wild Bootstrap	Paired Bootstrap	GEE
Incentive	0.368**	0.368**	0.368**	0.368**
	(0.025)	(0.054)	(0.025)	(0.025)
1 month post-intervention	0.117**	0.117**	0.117**	0.117**
	(0.023)	(0.016)	(0.023)	(0.022)
2 month post-intervention	0.103**	0.103**	0.103**	0.103**
	(0.023)	(0.033)	(0.024)	(0.023)
Fiveweek*Incentive	0.042	0.042	0.042	0.042
	(0.043)	(0.045)	(0.046)	(0.043)
Fiveweek*Month1	0.087†	0.087†	0.087†	0.087*
	(0.044)	(0.050)	(0.046)	(0.043)
Fiveweek*Month2	0.044	0.044	0.044	0.044
	(0.039)	(0.043)	(0.041)	(0.038)
Observations	403,922	403,922	403,922	403,922

Notes: Each column corresponds to the same regression as the first column in Table 4. The outcome variable is an indicator for whether or not the child ate at least one serving of fruits or vegetables. Each regression includes controls for gender, grade, day of week, and school fixed effects. Standard errors are clustered at the school level. **, *, and † signify statistical significance at the 1%, 5%, and 10% levels respectively.