Elisa DE MARCHI, Alessia CAVALIERE*, Rodolfo M. NAYGA, Alessandro BANTERLE

Incentivizing vegetable consumption in school-aged children: Evidence from a field experiment

Elisa De Marchi Post-Doc Reseracher, Department of Environmental Science and Policy, University of Milan, Italy, elisa.demarchi@unimi.it; Alessia Cavaliere* Assistant professor, Department of Environmental Science and Policy, University of Milan, Italy, alessia.cavaliere@unimi.it; Rodolfo M. Nayga, Jr. Distinguished Professor and Tyson Chair in Food Policy Economics, Department of Agricultural Economics and Agribusiness, University of Arkansas, Fayetteville, USA, rnayga@uark.edu; Alessandro Banterle, Full Professor, Department of Environmental Science and Policy, University of Milan, Italy, alessandro.banterle@unimi.it

ABSTRACT

We conducted a field experiment to test the effect of non-monetary incentives in increasing children's vegetable consumption during lunch at school. We measured children's daily vegetable consumption for 4 consecutive weeks prior to the provision of incentives, for 4 consecutive weeks during the incentive provision, and for 3 consecutive weeks right after the provision of incentives. To check the longer term effect of the incentive provision, we measured children's daily vegetable consumption 11 weeks after the post-intervention period. Results suggest that the incentives are effective in increasing vegetable consumption and that this effect persisted several weeks after the provision of the incentives ended. This is an important topic since gaining a better understanding of effects of non-monetary incentives can help in the design of nutrition and health policies aimed at improving the dietary behavior of children and potentially reducing childhood obesity.

Acknowledgments: The authors thank the children and teachers of the school (Educandato Statale Emanuela Setti Carraro dalla Chiesa, Milan) for participation in this project. The Ethical Committee of University of Milan approved the study.

INTRODUCTION

The rising rates of childhood obesity observed in many countries generally highlight the significant worsening of children's dietary quality (Grainger, Senauer, and Runge 2007; Raju, Rajagopal, Gilbride 2010; Hoy and Childers 2012). The nutritional quality of children's diet is important not only because it can have significant health consequences (i.e., increased disease risk) (WHO 2009), but also because it can hamper cognitive development and educational achievements (e.g., scholastic performance) (Lambert et al. 2004; Belot and James 2011; Hoy and Childers 2012; Ishdorj, Jensen, and Crepinsek 2013; Black, Johnston, and Peeters 2015). Furthermore, food consumption patterns during childhood play a crucial role in determining the wellbeing of individuals in the long run given that eating habits are developed at an early age and tend to persist throughout adulthood (Smith and Tasnadi 2007; Raju, Rajagopal, and Gilbride 2010).

One of the major contributors to poor nutrition among children is the low consumption of fruit and vegetables (FV) (Loewenstein, Price, and Volpp 2016), particularly for neophobic children (Birch and Fisher 1998; Galloway, Lee, and Birch 2003; Laureati, Bergamaschi and Pagliarini 2015). The inadequate consumption of FV has been the center of several information campaigns and school-based interventions that have been carried out in many countries in the past years (popular examples are represented by the '5-a-day' and the 'Feed Me Better' in UK, or the 'Making It Happen! School Nutrition Success Stories' promoted by USDA). However, the effectiveness of these campaigns in actually changing behaviors has been questioned (Belot, James, and Nolen 2016; Robertson 2008) and so finding ways to increase FV consumption among children is still an open challenge.

This issue is particularly relevant in the context of school cafeterias since children spend a significant portion of their day in the school environment, where many of them consume at least one main meal per day. Therefore, an inadequate consumption of FV at

school can compromise the overall balance of children's daily food intake (Grainger, Senauer, and Runge 2007; Caldeira et al. 2017). Recently, behavioral economists have begun to take important steps in this field of research by testing school-based interventions involving the provision of incentives to children (Loewenstein, Price, and Volpp 2016; List and Samek 2015; Belot, James, and Nolen 2016; Just and Price 2013). While the effectiveness of incentives in leveraging various behaviors has been acknowledged in several context for adults (such as smoking behavior, weight loss and cognitive task completion) (Shofield et al. 2015; Volpp et al. 2008; Volpp et al. 2009), only a few papers have focused on testing incentives among children (e.g., Just and Price 2013; List and Samek 2015; Belot, James, and Nolen 2016; Loewenstein, Price, and Volpp 2016). Perhaps, this is because the use of incentives with children raises some concerns. The main issue is related to the 'crowding out' effect; i.e., that incentives may reduce individual intrinsic motivation to complete a task or to undertake a specific behavior (Kamenica 2011; Gneezy, Meier, and Rey-Biel 2011). As such, there is concern that using incentives to nudge children towards various positive behaviors may, ultimately, backfire. Moreover, in the specific context of food behaviors, Newman and Taylor (1992) suggested that incentivizing food consumption may result in a decreased preference for those specific food items. Despite these concerns, several studies have demonstrated that the use of incentives can be effective in improving positive behaviors in children leading, in some cases, to habit formation. However, this line of research is relatively new and the results obtained up to now are still mixed.

To this purpose, we conducted a pilot study in Italy to examine the effectiveness of non-monetary incentive provision in increasing vegetable consumption among children in the elementary and the middle school. We collected data on children's vegetable consumption (i.e., using weight of vegetable leftover from lunch) before, during, and right after the incentive provision. Moreover, to further verify whether the effect of the incentives persists

long after the provision of incentives, we again measured children's vegetable consumption 11 weeks after the end of the post-intervention stage of the study. The aims of our study are to: (*i)* explore to what extent children in different age classes respond to non-monetary incentives, *(ii)* test the longer-term effect of incentive provision on vegetable consumption, and (*iii)* investigate the possible moderating role of children's' food neophobia in influencing children's vegetable consumption.

This paper contributes to the literature in several ways. Firstly, the results contribute in resolving prior conflicts in the literature on the effectiveness of non-monetary incentives in increasing vegetable consumption of schoolchildren. Secondly, by adopting a longer experimental period and follow-up, we provide novel insights on the effectiveness of incentives in creating longer lasting healthful habits. Thirdly, this pilot study extends current knowledge on the role of children's food neophobia and age in affecting vegetable consumption of children, two aspects that have been scarcely investigated yet. Moreover, to our best knowledge, this is the first time that an incentive-based study on school-aged children is conducted in a Mediterranean country. Italy has been historically characterized by the Mediterranean Diet model notoriously based on high consumption of FV. However, recent data show that eating patterns are rapidly evolving and the share of adult Italians meeting the WHO recommended daily amount of FV has dramatically declined over the last decade, particularly among the more disadvantaged population segments (Cavaliere, De Marchi and Banterle 2018; Cavaliere et al. 2019; Bonanno et al. 2017). This trend is likely to impact negatively on the food habits of the younger generations. Since this is one of the first studies of its kind conducted in a Mediterranean country, the results should provide a gauge on whether non-monetary incentives will work in different contexts, given that many of the past studies were conducted in the US. Additionally, Italy represents an interesting case study due to its peculiar situation in terms of obesity distribution (Banterle and Cavaliere 2014). While

this country still boasts one of the lowest adult obesity rates (10.3%) in Europe, the situation is opposite for children. According to the most recent available data (2015-2016) from the WHO Childhood Obesity Surveillance Initiative (COSI), Italy has one of the highest rates of childhood obesity in the EU, together with Cyprus in Greece, Malta, San Marino and Spain, where 1 in 5 boys (approximately 20%) are obese. These data are alarming, above all when compared to the much lower prevalence of childhood obesity (i.e., 5% to 9%) in Denmark, France, Ireland, Latvia, and Norway.

In this context, the results can provide some guidance for the design of school-based interventions geared at improving children diet quality at school, as well as in the formulation and implementation of future nutrition and health policies aimed at reducing childhood obesity on a larger scale.

This paper is organized as follows: the next section provides the relevant literature on which our experimental design is grounded; section 3 describes in detail the experimental design and procedures adopted; section 4 reports the main findings of our experiment by discussing the descriptive statistics and the econometric analysis; finally, section 5 provides the discussion of the results and the related conclusions.

LITERATURE BACKGROUND

Many of the past studies examining the effect of incentives on children's FV consumption were conducted in the US. Raju et al. (2010) in a large-scale study involving schoolchildren demonstrated the effectiveness of incentives, pledge and competition interventions in increasing healthy food choices at school, resulting in a significant increase in FV consumption which persisted after the intervention. This study also showed that children respond differently to the interventions based on their age, likely due to their different cognitive development stage, with younger children being more responsive to incentives

compared to the older ones. Just and Price (2013) conducted a field experiment in Utah to explore the impact of different incentive schemes (i.e., monetary incentives and lottery, either immediately provided or delayed) on elementary school children. They left the intervention in place for five days and showed that the incentives raised FV consumption overall. However, they found no evidence of medium-run effects, with consumption rates going back to baseline levels four weeks after the incentives were removed. List and Samek (2015) implemented another incentive-based intervention in after-school programs called Kids Café in the Chicago area, that provide free meals to low-income children and adolescents (6 to 18 years old). They used small prizes to encourage students to choose dried fruit over cookies, comparing different treatment schemes (positive *vs* negative framing, incentive alone, incentive paired with educational messages, educational messages alone). Furthermore, they compared the effect of short term (i.e., 1 day) *vs* long term (i.e., 5 days) interventions and found that incentives remarkably increased healthy choices, regardless of the type of framing both in the shortterm and long-term conditions. They also observed that the positive effect of incentives was sustained one week after the removal of the incentives, especially when the incentives were combined with the educational message. In a similar experiment conducted in Utah on elementary school children, Loewenstein, Price, and Volpp (2016) used financial rewards to encourage children to eat more FV at lunch. They implemented longer intervention periods (namely 3 weeks or 5 weeks) to further explore the power of incentives in leading to habit formation. Their results confirmed the effectiveness of incentives in increasing FV consumption and showed that the effect of the intervention persisted in the long run.

Specifically, the authors observed that two-months after the incentives were removed, FV consumption rates remained considerably higher than the baseline. Furthermore, they observed that the long-run effect was stronger (although marginally) for children in the 5-

weeks condition, which seems to indicate that longer interventions may produce more persistent response.

Positive results were also obtained by Belot and James (2016) which used non-financial incentives to encourage children to choose more FV during lunchtime at school. Their field experiment was conducted in England and involved children from the second and the fifth grades. They observed increased FV consumption rates during the incentive period. However, they found only little evidence of persistence of the effects six months after the incentives were removed.

Overall, the findings from past studies with regard to the effectiveness of incentives on habit formation are quite mixed. This may be, at least in part, attributable to the fact that there is no common understanding yet on the psychological mechanisms underlying habit formation (Neal et al. 2011). It is known that habits are formed when individuals develop implicit associations in memory between contexts and responses when repeatedly undertaking a specific action (Neal et al. 2011; Carden and Wood 2018). Most researchers believe that attentional mechanisms, including the use of incentives, play a crucial role. Indeed, people tend to repeat actions that are rewarding or motivated by specific goals. In other words, rewarding actions gain an attentional priority over non-rewarding ones, which facilitates implicit association in memory and thus, habit formation (Neal et al. 2011; Carden and Wood 2018). However, this is not always the case. Previous findings have shown that habits can be merely activated by context cues, with very little influence of goals (e.g., Neal et al. 2011) and that rewards may lead individuals to deliberate about the repeated behavior, ultimately precluding habit formation. Furthermore there is still uncertainty regarding habit strength and persistence, which are affected by frequency and contexts (Gillan et al. 2015).

Such inconsistent evidence further motivates our study. Additionally, as previously mentioned, we decided to account for children's food neophobia as a personality trait since

this could potentially influence vegetable consumption (Pliner and Hobden 1992). Food neophobia can be described as the 'fear of new food', which results in a reluctance to eat novel and unknown food items, thus having a strong impact on the overall diet quality and on the development of individual food preferences (Birch and Fisher 1998; Pliner and Hobden 1992; Russell and Worsley 2008). Most importantly, food neophobia is proven to be associated with low consumption of FV (Cooke, Carnell, and Wardle 2006; Nicklaus et al. 2005; Galloway, Lee, and Birch 2003). Furthermore, even though food neophobia can be observed at all stages in life, previous studies suggest that it is particularly evident during childhood, thus having a crucial role in taste development and food habit formation (Cooke, Carnell, and Wardle 2006; Pliner and Salvy 2006; Galloway, Lee, and Birch 2003). We discuss how we measured food neophobia in more detail in section 3.4.

METHODS

Experimental Design

To analyze the effectiveness of non-monetary incentives in increasing vegetable[1] consumption, eventually leading to positive habit formation, we conducted a field experiment in Milan (northern Italy) in a public school that includes both elementary school grade levels (1st to 5th grade, 6 to 11 years old) and middle school grade levels (6th to 9th grade, 11 to 14 years old). Data were collected over a period of 11 weeks between September and December 2017, while the follow-up study (one week long) was conducted in February 2018. Before the beginning of the school year, all teachers and the canteen operators were informed about the study, its main scope, and all the procedures to be adopted. Parents were also informed about the project through a detailed letter delivered in person to them by the teachers and published in the parents' private area of the school website. Parents were explicitly asked not

1. In this paper we decided to focus exclusively on vegetable consumption instead of considering also fruits, due to the fact that in our school fruit is always served in class during the morning break.

to reveal to their children any information about the project in order to avoid influencing their behavior. Parents could freely decide whether to have data collected about their child by signing a consent form or not. After collecting the parents' permissions, we divided the school children into two experimental groups: the intervention group that received the non-monetary incentives, and the control group. Assignment to one of the two experimental groups was made randomly and randomization was done at the classroom level as follows.

The elementary grade levels have 10 classes: two 1st grade classes, two 2nd grade classes, two 3rd grade, two 4th grade, and two 5th grade classes. Similarly, children in the middle grade levels are distributed into 6 classes, namely, two 6th grade, two 7th grade, and two 8th grade classes. Each class has 24 students, on average. Being interested in exploring age differences, we were careful that age classes were evenly distributed in each experimental group. Accordingly, one class in each grade was randomly assigned to the control group or the intervention group. Randomization in terms of socio-economic and demographic characteristics within and between the two groups was guaranteed *ex-ante* by means of the school enrollment procedure. Specifically, the children in Italy are enrolled in public schools depending on how close they live to the school. They are then randomly assigned to each class according to their age. Gender is roughly balanced in each class and no other selection criteria are adopted. In order to test the effectiveness of the incentive provision in increasing vegetable consumption at lunch and to test whether this leads to positive habit formation, we randomly assigned classes in each grade level to either the control group or the treated group and used the weight of the vegetable leftovers at lunch as our outcome measure. Such outcome measure was adopted following the approach used in previous studies (van Kleef et al. 2014; van Kleef, Bruggers, and de Vet 2015). Baseline data were collected for each class for 4 weeks (T_0) without children in both groups being aware of the experiment being conducted. In the following 4 weeks (T_1), children in the intervention group received the incentives, while

227 no changes occurred for the control group. In the next 3 weeks (T_2) after the intervention
228 period, the incentives were removed from the treated group and the experimenters continued
229 the monitoring of vegetable leftovers of both experimental groups. Additionally, to verify the
230 long-term effect, we conducted a one-week follow-up study 11 weeks after the end of the
231 post-intervention period and monitored vegetable consumption of both experimental groups
232 (T_3).
233

234 Experimental Procedure

235 The school provides students a daily menu with a first dish and a second dish always
236 including a portion of vegetables composed of salad and cooked vegetables. The cooked
237 vegetables include five varieties that vary daily (spinach, green beans, baby carrots, broccoli,
238 and a vegetable mix with potatoes, carrots and zucchini). In order to obtain accurate
239 measures of vegetables leftovers, the canteen operators were instructed to serve roughly
240 equal portions using specified spoons. For the first week, before the service started,
241 experimenters weighed three vegetable portions (salad + daily cooked vegetable) and the
242 average weight was used as reference point to quantify the leftovers. After lunch, the
243 experimenters collected the vegetable leftovers of each class in separate transparent plastic
244 bags and weighed the content[2]. In the lunchroom, each class is usually assigned to specific
245 tables, which helped the experimenters to separately collect leftovers without making any
246 change in the cafeteria environment.

247 Although we really wanted to collect our outcome measure (i.e., vegetable leftovers) at
248 the individual student level, we were only able to do this at the class level. We asked for
249 permission to collect individual leftover data in order to be able to detect differences in

2. So as not to generate concerns among children seeing their vegetable leftovers being collected, the experimenters were dressed as canteen operators and the leftover collection was conducted while the tables were being cleaned following the usual procedure.

250 vegetable consumption based on personal characteristics, but we were not allowed to do so
251 because of several restrictions imposed both by the experimental setting and the Ethical
252 Committee that evaluated and approved our research project[3]. Furthermore, the children only
253 have 30 minutes for their lunch and the lunchroom has to be cleaned and set up very quickly
254 before the arrival of the second shift (the high school students) and so this has severely
255 limited the time available to us. This made it impossible to separately collect individual
256 leftovers, as this procedure would have taken a longer time. Also our request to use small
257 cameras in the cafeteria was denied by the Ethics Committee (i.e., counterpart of IRB in the US
258 system) for privacy reasons related to the young age of the subjects involved. Before the start
259 of the experiment, the teachers responsible for the classes in the intervention group were
260 provided with the instruction form to be read to the students just prior the start of the
261 incentive period. The instructions were differentiated across grades in order to adapt the
262 language to the different age categories of the children. Children were told that they would
263 receive a prize at the end of the week if they finish the whole portion of vegetables served
264 each day to them from Monday to Friday, including both salad and the cooked vegetable. In
265 order to avoid contamination across the two experimental groups, children were instructed to
266 keep the experiment secret and not to share any information with peers in other classes, or
267 else they will face the penalty of the exclusion of the whole class from the incentive provision.
268 Furthermore, to avoid cheating behavior aimed at receiving the prize, the children
269 were told about the importance of being honest and were informed that throwing vegetables
270 on the floor, sharing/changing plates with peers, or hiding vegetables would result in the
271 exclusion of the entire class from the experiment. Children in the intervention group were
272 also told that no additional prizes would be given for additional portions of vegetables eaten.
273 These rules were repeated to them three times during the first intervention week and then

3. All approval documents are available upon request.

each Monday for the following three weeks. The identification of the children who deserved to receive the incentive was made possible through the use of small tags with the children's names placed on the trays by the teachers[4]. The use of these tags was introduced both for children in the control and in the incentive condition to avoid any perceived difference between the two groups. Tags with the names were introduced at the very beginning of the school year (namely, two weeks before the beginning of the experiment) and presented to children as a novel rule adopted by the school to facilitate the familiarization of the teachers with the students' names. Incentives were distributed by the teachers on Friday of each week, right before the children go home. This procedure was aimed at limiting contamination between the two experimental groups. Children with special dietary requirements participated in the study although their consumption data were excluded from the analysis[5]. At the end of the experiment, all children were asked to complete a questionnaire to collect information on food neophobia and vegetable liking (Appendix A). The way the questions were asked and the response scales used were differentiated based on the age categories of the children.

Incentives: Type and Timing

In this study, we used non-monetary incentives for several reasons. Firstly, the schools generally prefer this type of rewards (Levitt et al. 2016). Non-monetary prizes are commonly used to reward children for winning school competitions (e.g., math championship). Hence, the teachers as well as the children are comfortable with the use of these prizes. This familiarity makes them potentially more responsive to non-monetary incentives than to cash-based rewards (Levitt et al. 2016). This is a crucial aspect especially for younger children who

4. Only the first name was indicated on the tag, but no surnames, for privacy reasons imposed by the Ethical Committee

5. Special diets are pre-prepared and served in different plates. This allowed the experimenters to easily recognize them on each table.

may be unable to fully understand yet the real value of money (Levitt et al. 2016; Just and Price, 2013). A second motivation for the use of non-monetary rewards is related to the lower financial commitment needed to implement such intervention. The economic value of the prizes used in previous studies is generally less than one dollar (Belot, James, and Nolen 2016; Levitt et al. 2016; List and Samek 2015; Just and Price 2013). The fact that all of these studies obtained significant positive results despite the small amount of the monetary value of the rewards represents a key point in terms of replicability of the study and policy implication.

The low cost-effectiveness of the incentives may encourage policy makers to evaluate the implementation of school-based interventions of this type on a large scale. In contrast with previous studies that used the same type of incentive both for children and adolescents (List and Samek 2015), we targeted and tested our incentives for different age groups. For the elementary grades, we selected different prize options whose value ranged between 0.40 and 0.80 euros. We pre-tested these incentives to a small sample of 35 elementary school children from another public school in Milan (1st to 5th grade) by asking them to respond to a brief questionnaire, which consisted of the images of eight different prizes. Children were asked to mark their top 4 favorite options among the 8 presented. We then selected the four items that obtained the highest scores (Table 1). As for the middle school, we followed the same approach, but we proposed prizes with a slightly higher economic value (between 0.85 euro and 1.15 euros). Similar to what was done for the elementary grades, we pre-tested the incentives by asking 31 middle school children from a different public school (6th to 8th grade) to rate all the nine items presented on a 5-points Likert scale where low values corresponded to low liking and high values represented high liking (Table 1). The four prizes with the highest scores were selected as incentives (Appendix A). This procedure allowed us to have incentives that were salient enough for the children in different age groups. As for how long it takes to form longer term habits, previous studies provided mixed results. For instance, List

322 and Samek (2015) implemented 1-day or 5-day interventions and found that the effect of
323 incentives was sustained one week after the reward removal. Just and Price (2013), on the
324 other hand, demonstrated that five days may not be sufficient enough to determine habit
325 formation. Loewenstein, Price, and Volpp (2016) tested the effectiveness of longer
326 intervention periods (3-week and 5-weeks). Their results suggested that the longer the
327 intervention period, the stronger the effect of the incentives. Furthermore, they found that the
328 effect of the intervention persisted even two months after the incentive provision. Belot,
329 James, and Nolen (2016) also obtained positive results with a 4-week intervention period,
330 even though they did not find longer term effects six months later. Overall, there is high
331 variability in terms of the time required to form a habit, which depends both on individual
332 characteristics and on the specific behavior involved (Gardner, Lally, and Wardle 2012; Lally
333 et al. 2010). As such, and given that it is not possible to precisely establish how long it would
334 take to form a positive consumption habit in children, we decided to follow the conventional
335 wisdom that habit formation should occur in 21 days (Gardner, Lally, and Wardle 2012;
336 Loewenstein, Price, and Volpp 2016).

337 PLEASE INSERT TABLE 1 ABOUT HERE

338

339 Food Neophobia

340 Food neophobia has been found to remarkably affect the variety of foods that children
341 are inclined to taste and eat and to be related to lower consumption of FV (Cooke, Carnell, and
342 Wardle 2006; Nicklaus et al. 2005; Galloway, Lee, and Birch 2003). A well-known scale to
343 measure food neophobia in children (The Child Food Neophobia Scale - CFNS) has been
344 proposed by Pliner in 1994. The scale was based on the adaptation of the items of the adult
345 food neophobia scale, with the aim of capturing children's behavior. The main limitation of
346 this scale is that it implies that responses are given by parents. In other words, children's food

347 neophobia elicited through the CFNS is actually based on parents' evaluation of how inclined
348 their children are to try novel foods, which may lead to over- or under-estimation biases
349 (Aldridge, Dovey, and Halford 2009). To overcome this issue, we used a scale based on
350 children's self-reported responses: the Italian Child Food Neophobia Scale (ICFNS). The scale
351 was validated by Laureati, Bergamaschi, and Pagliarini (2015) on a large sample of Italian
352 primary school children. The ICFNS consists of 8 items (4 neophobic and 4 neophilic) phrased
353 with a simple age-appropriate vocabulary describing food consumption contexts likely to be
354 familiar to children. Responses are based on a 5 point scale corresponding to the following
355 five statements: 'Very false for me', 'False for me', 'So-so', 'True for me', and 'Very true for me'
356 (Laureati, Bergamaschi, and Pagliarini 2015). The main advantage, besides being validated on
357 Italian children, is that each value is paired with a facial expression that helps children,
358 especially those in lower grades, interpret the numeric values, making it possible to obtain
359 self-reported data. As for students in the 6th, 7th and 8th grades, we used the same items, but
360 removing faces from the response scale accounting for age differences.
361

362 **RESULTS**

363 Descriptive Analysis

364 Our sample consisted of 370 children. As for gender, females were overrepresented
365 compared to males (60% females and 40% males) but the gender distribution was roughly
366 homogeneous across the experimental conditions (109 females and 79 males in the control
367 group; 113 females and 69 males in the incentive group). As a first step, we analyzed some
368 baseline characteristics of the children in the two experimental conditions based on the
369 information obtained from the questionnaire responses. Table 2 exhibits the p-values for the
370 t-test on children's liking of the vegetables served for lunch, including salad and cooked
371 vegetables served with daily rotation, namely green beans, broccoli, vegetable mix, carrots,

372 and spinach. Children were shown pictures of these vegetables and were then asked to rate
373 their degree of liking using a 7-point scale. As for the younger children (1st to 5th grades), we
374 followed the approach validated by Pagliarini, Gabbiadini, and Ratti (2005), where the 7-point
375 numeric scale was paired with a hedonic-facial scale to help them interpret the numeric
376 values. We found no significant differences across the control and the incentive groups.
377 Similar results emerged when analyzing children's liking by vegetable type, except for carrots,
378 which were slightly preferred by children in the control condition. We also checked for
379 differences in children's level of food neophobia across the two experimental conditions. The
380 individual responses to the ICFNS were used to obtain class-level mean values, which are
381 reported in Table 3. Also in this case, we did not find significant differences between children
382 in the control and in the incentive group (Table 2). Overall, these statistics suggest that the
383 randomization was successful in balancing the observable characteristics of the two
384 experimental groups.

385 PLEASE INSERT TABLE 2 AND 3 ABOUT HERE

386 The same pattern of non-significant differences between groups can be observed when
387 analyzing baseline (pre-intervention) vegetable leftovers, weighed daily at the class level.
388 Table 4 reports the p-value of the t-test for whether the baseline vegetable leftovers differ
389 between the control and the incentive groups. At T_0 (from week 1 to week 4), daily leftovers of
390 the classes in the control amount to almost 659.9 grams, which is not statistically significantly
391 different from the 633.3 grams of daily leftovers from the classes in the treated group.

392 PLEASE INSERT TABLE 4 ABOUT HERE

393 The mean leftover values indicate that during the pre-intervention period, children in
394 both groups ate on average 60.4% of the served daily vegetable portion. We cannot establish
395 whether this value is in line with past studies since data are quite mixed. Just and Price (2013)
396 for instance registered baseline consumption rates of fruit and vegetables around 33%, while

Loewenstein, Price, and Volpp (2016) measured them at around 40%. Belot, James, and Nolen (2016) reported much higher baseline consumption, around 76% for their overall sample. This high variability may be due to a number of factors, including for instance country-specific food patterns and traditions. We then analyzed differences in vegetable leftovers at T_1, T_2 and T_3 to examine the effect of the incentive provision in reducing leftovers of the treated group.

As for T_1 (from week 5 to week 8), the p-value (0.001) of the effect size suggests that the incentives were effective in reducing vegetable leftovers of the treated classes compared to their peers in the control group in the same period. We found significant differences between the two groups also at T_2 (p-value 0.008), that is after the incentives were removed. In fact, vegetable leftovers of the incentive group remained lower than those of the control group suggesting that the effect of incentives persisted after the end of the treatment. Remarkably, we obtained the same significant result also at T_3, namely two months after the end of the experiment (that is, 11 weeks after the incentives were removed). The follow up data collected at T_3 indicate that leftovers of the incentive group remained lower than the baseline and lower relative to the leftovers of children in the control group. Moreover, vegetable leftovers of the control group did not vary significantly overall in all periods, except for a slight increase from T_0 to T_1. To explore this issue further, we conducted a paired t-test using the leftover data of the control group at T_0 and T_1, and confirmed that there is no statistically significant difference in leftover means between the two experimental periods. (t_{91}= 1.583, p <0.117). Despite the extensive procedures we applied to avoid contamination across the two experimental groups, as previously discussed, it is still possible that this slight increase in vegetable leftovers of the control group from T_0 to T_1, although not significant, could be indicative of some level of contamination. For example, it is possible that some children in the control group who may have heard that someone else received a reward for eating vegetables might be inclined to eat fewer vegetables out of spite. However, the increase

422 in vegetable waste by the control group seems to go away once the incentives are no longer in
423 place. Indeed, vegetable leftovers of the incentive group in T_1 are significantly lower than that
424 of the control group during the baseline period T_0 (i.e., 473.6 vs 659.9) and we found
425 statistically significant differences in vegetable leftovers between the two groups after the
426 intervention in periods T_2 and T_3 despite the decrease in leftovers in the control group after
427 period T_1[6].

428 A graphic representation of the vegetable leftover patterns of the control and the
429 incentive groups is illustrated in the graphs of Figures 1 and 2. Figure 1 illustrates the overall
430 leftover pattern of two experimental groups over the 11 experimental weeks and the follow
431 up. To analyze more in detail variations in vegetable leftovers of the groups we split the
432 graph in Figure 1 into three separate graphs respectively corresponding to T_0, T_1 and T_2
433 (Figure 2). Graph A in Figure 2 reports baseline leftover data respectively for the children in
434 the control and the incentive condition. The curves are close to each other and follow the
435 same curve trend, which is in line with the t-test results described above that did not detect
436 significant differences. Vegetable leftovers at T_1 are represented in Graph B.

437 PLEASE INSERT FIGURES 1 AND 2 ABOUT HERE

438 While leftovers of children in the incentive group have decreased with respect to T_0,
439 the upper curve of the control group highlights the difference in leftovers between the two
440 conditions. Graph B also shows that leftovers of both groups followed the same fluctuating
441 pattern during the intervention weeks. They slightly decreased at the end of week 5,
442 increased during week 6, and decreased again during week 8. This may be attributable to the
443 fact that the vegetables, being prepared daily, may be subject to slight variations in taste.
444 Finally, Graph C illustrates the leftover patterns at T_2, namely when the incentives were not

6. We were not able to conduct the study in two different schools because almost all the schools in the city have meals provided by external suppliers who were unwilling to participate in the study mainly due to time constraints and limited number of employees. The school where we collected our data operated its own cafeteria and hence had the flexibility to help us with the data collection and conduct of the experiment.

provided anymore. Although the distance between the curves of the control and the incentive groups is less pronounced in comparison with Graph B, the leftovers of children who received the treatment remained lower than those of the control group and lower than the baseline measures.

Econometric Analysis

Table 5 reports the regression based results for whether the treatment was effective in reducing daily vegetable leftovers, which represents our outcome variable. We used the log of daily leftover weight of each class as our dependent variable. The leftovers of children with special dietary requirements were not considered in the analysis. To further analyze the treatment effect and its variation across the experimental periods, and to explore the role of children's age and/or vegetable variety as well as their neophobia, we estimated five models with different specifications. Model 1, the baseline model, only includes the treatment variable as a regressor.

Model 2 further explores the effect of the treatment over time by adding the interaction terms of the treatment respectively with T_1, T_2, and T_3 variables. Model 3 adds the grade fixed effects to Model 2 specification while Model 4 further adds the fixed effects for vegetable types. Finally, the neophobia variable is added in Model 5.

Results from Model 1 indicate that children in the treated group have lower vegetable leftovers (32% lower) than the children in the control group. This first evidence seems to support the effectiveness of non-monetary incentives in leveraging higher vegetable consumption in children. The second Model (Model 2) adds to the baseline the interaction terms between the treatment and T_1, T_2, and T_3. This allows us to explore how the treatment effect varies across the experimental periods. The results show that the treatment effect was more evident at T_1; that is when incentives were provided to the treated group. In this

experimental period, leftovers of the treated children reduced by almost 46% compared to vegetable leftovers of children in the control group. At T_2, when the incentives were removed, the leftover reduction is less than in T_1, as expected, but the leftovers of children in the incentive group remain lower than those of their counterparts in the control group. Interestingly, the significant interaction of the treatment variable with T_3 (i.e., the follow up period, 11 weeks after the experiment was concluded) highlights a further reduction of leftovers of the incentive group, with values comparable to those observed at T_1. Given that neither the vegetable varieties nor the way they were prepared varied across the 4 experimental periods, we can exclude the possibility that such reduction in vegetable leftovers was caused by variation in children's preferences or tastes. Admittedly, given that our T_3 period was only for one week, future studies should further explore the actual persistence of the occurred variation in consumption behavior.

Model 3 includes grade fixed effects to control for age. The parameter estimates for the grade dummy variables show that children in the 2nd and 6th grade eat less vegetables (i.e., have higher leftovers) compared to children in the 1st grade, whilst those in the 4th and 5th grade have significantly lower leftovers relative to their 6 year old peers. Overall, this seems to highlight that there is heterogeneity in behavior across age classes, with 9 and 10 year old children eating more vegetables compared to others.

Model 4 additionally controls for differences in leftovers by accounting for the different types of vegetables offered to children with daily variation, as mentioned in section 4.1. The coefficients suggest that leftovers of the vegetable mix and carrots are lower with respect to green beans, which is the preferred vegetable type. Spinach leftovers, however, are on average higher compared to those of green beans. No significant differences were found for broccoli. An important insight emerging from these models is that the treatment effect is robust to the addition of grade and vegetable type fixed effects. Indeed, in all models the coefficients of the

495 treatment and its interactions with T_1, T_2 and T_3 remain almost unvaried both in terms of
496 magnitude and significance.

497 PLEASE INSERT TABLE 5 ABOUT HERE

498 Moreover, we run an additional regression to explore the role of children's food
499 neophobia. Individual responses to the ICFNS were used to estimate the mean neophobia
500 score of each class which was then included in Model 5. The significant and positive
501 coefficient (0.566) indicates that high neophobia scores are associated with higher vegetable
502 leftovers. While this result does not obviously suggest causation, this effect seems to suggest
503 that neophobia is closely related with children's willingness to eat different types of
504 vegetables. Accordingly, it is reasonable to expect that neophobia may negatively influence
505 the treatment effectiveness. In other words, even though we cannot derive conclusions on this
506 aspect based on the present analysis, it is plausible that children with high neophobia respond
507 to incentives to a lesser extent compared to children with low neophobia levels.

508 Overall our findings highlight three main patterns: (i) providing non-monetary
509 incentives can increase the amount of vegetables eaten (i.e., reduced leftovers) by children (ii)
510 the effect of incentives seems to have a long-term positive impact on vegetable consumption,
511 which is sustained even 11 weeks after the rewards were removed; and (iii) high neophobia is
512 associated with higher vegetable leftovers.

513

514 **DISCUSSION AND CONCLUSIONS**

515 We carried out a pilot school-based field experiment in one school in Italy in order to
516 investigate to what extent children respond to non-monetary incentives and whether
517 incentive provision may result in positive longer-term effects on vegetable consumption. We
518 chose to test the effect of incentives since incentive-based interventions can be easily
519 implemented, require relatively modest financial investments, and are therefore suitable to be

520 used on a large scale (Raju, Rajagopal, and Gilbride 2010). In this study, we also explored the
521 role of food neophobia and children's age in influencing vegetable consumption.
522 Overall, the results of our pilot study suggest that the use of small incentives can
523 be successful in increasing children's vegetable consumption rates (Belot, James, and Nolen
524 2016; Loewenstein, Price, and Volpp 2016; List and Samek 2015; Just and Price 2013).
525 Differences in our outcome measure, that is daily vegetable leftover weight of each
526 class, seem to indicate that incentives could lead to a reduction in vegetable leftovers. During
527 the incentive period, we observed a 50% lower vegetable leftovers from children in the
528 treated group than children in the control group and this effect seemed to persist in the longer
529 run. Indeed, leftovers of the treated group remained about 13% lower than those of their
530 control counterpart during the three weeks that immediately followed the incentive period.
531 We also observed that the incentive effect was lasting 11 weeks post-intervention,
532 given that weighed leftovers of the treated group were about 40% lower than those of the
533 control group. As mentioned in the experimental procedure section, we were not allowed to
534 collect leftovers at the individual level. As such, we cannot establish whether the incentives
535 were effective for all children or solely to some pupils in each class. In other words, it could be
536 that only some children in each class responded to the incentives contributing to the observed
537 reduction in leftovers, with others not modifying their eating behavior at all. Nonetheless, the
538 results of our pilot study seem to suggest that the repetition of a behavior in a consistent
539 setting can lead to habit formation and that the attentional mechanism generated through
540 incentive provision may facilitate this process (Gardner, Lally, and Wardle 2012; Neal, Wu,
541 and Kurlander 2011; Lally et al. 2010).
542 Loewenstein, Price, and Volpp (2016) proposed two additional explanations upon
543 which incentives may result in habit formation. One is that incentives may encourage children
544 to eat food items that are usually avoided, thus making them discover novel tastes. In other

545 words, incentives may provide an extrinsic motivation to try novel foods, or to re-start eating
546 known items that were abandoned with no specific reason. The second mechanism is related
547 to social norms. It could be that incentives may contribute to making vegetable consumption
548 more popular among children, thus making it more appealing and leading to increased intake
549 (Loewenstein, Price, and Volpp 2016).

550 Another aspect that is worth mentioning is that we did not find evidence of any
551 crowding out effect. If crowding out occurred in our study, this would have been likely
552 reflected in vegetable leftovers of the incentive group going back to baseline levels or below.
553 Instead, our results, together with similar findings provided by past studies (i.e., Just and Price
554 2013; Loewenstein, Price, and Volpp 2016; Belot, James, and Nolen 2016), seem to reinforce
555 the evidence that incentives do not significantly affect the inner motivation of children.

556 Moreover, we were able to detect heterogeneous consumption across age classes, with
557 children in the 2nd and 6th grade eating less vegetables than those in the 1st grade, and children
558 in the 4th and 5th grades eating significantly more than their 6 year old peers. Prior research
559 on incentives provided limited insights on this specific issue. The paper by Raju, Rajagopal,
560 and Gilbride (2010) is one of the very few that examined this aspect. They found that younger
561 pupils responded more favorably to incentives compared to the older ones and ascribed such
562 differences to the level of cognitive development reached by children in each age class. It is
563 also interesting to note that no significant results were found with regard to grades 7th and 8th,
564 which suggests a boundary condition for the value of non-monetary incentives as children
565 age. Although our results suggest that age can influence the way children respond to
566 incentives, we cannot identify a specific pattern in our data that would allow us to derive
567 more robust conclusions.

568 Additionally, we observed that children's food neophobia is related to a decrease in
569 vegetable consumption, which is in line with previous studies (Birch and Fisher 1998;

570 Galloway, Lee, and Birch 2003; Laureati, Bergamaschi, and Pagliarini 2015). This indicates
571 that food neophobia may have a key role in predicting the amount of vegetables that children
572 eat, thus highlighting the importance of accounting for this personality trait in experimental
573 studies of this type. Indeed, given the relationship between neophobia and vegetable
574 consumption, it is reasonable to expect that children's fear of novel foods may ultimately
575 compromise the effectiveness of incentives in reducing vegetable leftovers. Furthermore, this
576 may have implications for the practical implementation of incentive schemes such as the one
577 tested in our pilot study and those of past studies since it suggests that more effort is likely
578 needed to change vegetable consumption habits of food neophobic children.

579 Overall, we believe that the results of this pilot study contribute to resolving prior
580 conflict in the literature that non-monetary incentives can increase healthful consumption in
581 young elementary age children. Our study also investigated group behavior by introducing a
582 new moderating variable that may explain some of the past inconsistent results, food
583 neophobia. By using a relatively longer study period, we were able to assess the longer term
584 effect of incentives in children and more broadly contribute to the habit formation literature.
585 We were also able to observe a wide age range of children and hence were able to detect that
586 our incentives may not work with older children or early teens. Finally, we were able to
587 expand the applicability of the use of incentives in school children by showing the positive
588 effects on vegetable consumption for children outside of the US where most of the current
589 literature is based.

590 Our findings however need to be taken with caution given that we were limited in our
591 ability to collect individual level data. We were also only able to collect data from a single
592 school due to the challenges of getting other schools in our study area to cooperate. This is not
593 ideal because the experimental procedure we followed does not allow us to exclude with
594 certainty that some contamination occurred between the control and the intervention groups.

Nevertheless, we were particularly careful about this aspect and we properly instructed all people involved in the study as well as the children not to share this experience with others. We cannot definitively rule out however the possibility that someone failed to maintain the vail of silence.

Given these limitations, future research should attempt to collect individual level data from schoolchildren to test the robustness of our findings. To further extend knowledge on this topic, future studies should also account for some factors related to children' eating behaviors that have not been accounted for in our study. Parental influence, for instance, has been demonstrated to shape children eating behaviors in various ways (Yu 2011; Hoi and Childres 2012) and so future research that would take this into account in relation to children's dietary behavior would make a substantial contribution to the understanding of the effectiveness of incentives.

REFERENCES

Aldridge, Victoria, Terence Dovey, and Jason C. G. Halford. 2009. The role of familiarity in dietary development. *Developmental Review,* 29: 32-44.

Banterle, Alessandro and Alessia Cavaliere. 2014. Is there a relationship between product attributes, nutrition labels and excess weight? Evidence from an Italian region. *Food Policy*, 49: 241-249.

Belot, Michèle and Jonathan James. 2011. Healthy school meals and educational outcomes. *Journal of Health Economics,* 30: 489-504.

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

619 Birch, Leann L., and Jennifer Fisher. 1998. Development of eating behaviors among children
620 and adolescents. *Pediatrics,* 101: 539-549.
621 Black, Nicole, David W. Johnston, and Anna Peeters. 2015. Childhood Obesity and Cognitive
622 Achievement. *Health Economics,* 24: 1082-1100.
623 Bonanno, Alessandro, Francesco Bimbo, Elena Castellari, and Paolo Sckokai. 2017. Five-a-Day,
624 Fruit and Vegetables Portions, and the Food Environment: The Italian Case. *Applied*
625 *Economic Perspectives and Policy,* 39: 682- 709.
626 Caldeira, Sandra, Stefan S. G. Bonsmann, Ioanna Bakogianni, Charmaine Gauci, Antoinette
627 Calleja, and Artur Furtado. 2017. Public Procurement of Food for Health. *Technical*
628 *report on the school setting*, 1-88.
629 Carden, Lucas, and Wendy Wood. 2018. Habit formation and change. *Current Opinion in*
630 *Behavioral Sciences,* 20: 117–122.
631 Cavaliere, Alessia, Elisa De Marchi, and Alessandro Banterle. 2018. Exploring the adherence to
632 the Mediterranean diet and its relationship with individual lifestyle: The role of healthy
633 behaviors, pro-environmental behaviors, income, and education. *Nutrients,* 10: 141.
634 Cavaliere, Alessia, Elisa De Marchi, Franco Donzelli, and Alessandro Banterle. 2019. Is the
635 Mediterranean Diet for all? An analysis of socioeconomic inequalities and food
636 consumption in Italy. *British Food Journal,* forthcoming.
637 Cooke, Lucy J., Susan Carnell, and Jane Wardle. 2006. Food neophobia and mealtime food
638 consumption in 4–5 year old children. International Journal of Behavioral. *Nutrition*
639 *and Physical Activity,* 6: 3-14.
640 Galloway, Amy T., Yoonna Lee, and Leann L. Birch. 2003. Predictors and consequences of food
641 neophobia and pickiness in young girls. *Journal of American Dietetic Association,* 103:
642 692-698.

Gardner, Benjamin, Phillippa Lally, and Jane Wardle. 2012. Making health habitual: the psychology of "habit-formation" and general practice. *The British Journal of General Practice,* 62: 664-666.

Gillan, C.M., Otto, A.R., Phelps, E.A., and Daw, N.D. 2015. Model-based learning protects against forming habits. *Cognitive Affective and Behavioral Neuroscience*, 15: 523-536.

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. When and Why Incentives (Don't) Work to Modify Behavior. *The Journal of Economic Perspectives,* 25: 191-209.

Grainger, Corbett, Senaure, Banjamin, and C.Ford Runge. 2007. Nutritional Improvements and Student Food Choices in a School Lunch Program. *Journal of Consumer Affairs*, 41(2): 265-284.

Hoy, Mariea G., and Courtney Childres. 2012. Trends in Food Attitudes and Behaviors Among Adults with 6–11-Year-Old Children*. Journal of Consumer Affairs*, 43(6): 556-572.

Ishdorj, Ariun, Mary Kay, Crepinsek, and Helen H. Jensen. 2013. Children's Consumption of Fruits and Vegetables: Do School Environment and Policies Affect Choices at School and Away from School? *Applied Economic Perspectives and Policy,* 35: 341-359.

Jalava, Nina, Juanna S. Joensen, and Elin Pellas. 2015. Grades and rank: Impacts of non-financial incentives on test performance. *Journal of Economic Behavior & Organization,* 115: 161-196.

Just, David R., and Joseph Price. 2013. Using incentives to encourage healthy eating in children. *Journal of Human Resources,* 48: 885-872.

Kamenica, Emir 2011. Behavioral Economics and Psychology of Incentives. *Annual Review of Economics,* 4: 427-452.

Lambert, Janet, Carlo Agostoni, Ibrahim Elmadfa, Karin Hulsof, Edburga Krause, Barbara Livingstone, Piotr Socha, Daphne Pannemans, and Sonia Samartins. 2004. Dietary

intake and nutritional status of children and adolescents in Europe. *British Journal of Nutrition,* 92: S147–211.

Lally, Phillippa, Cornelia H. M. Van Jaarsveld, Henry W. Potts, and Jane Wardle. 2010. How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology,* 40: 998-1009.

Laureati, Monica, Valentina Bergamaschi, and Ella Pagliarini. 2015. Assessing childhood food neophobia: Validation of a scale in Italian primary school children. *Food Quality and Preference* 40: 8-15.

Levitt, Steven D., John A. List, Susanne Neckermann, and Sally Sadoff. 2016. The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance. *American Economic Journal: Economic Policy,* 8: 183-219.

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

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

Nardone Paola. 2016. OKkio alla salute: risultati 2016 sugli stili di vita dei bambini.

Neal, D. T., Wood, W., Wu, M., and Kurlander, D. 2011. The pull of the past: When do habits persist despite conflict with motives? *Personality and Social Psychology Bulletin,* 37, 1428–1437.

Newman, Joan, and Alan Taylor. 1992. Effect of a means-end contingency on young children's food preferences. *Journal of Experimental Child Psychology,* 53: 200-216.

Nicklaus, Sophie, Vincent Boggio, Claire Chabanet, and Sylvie Issanchou. 2005. A prospective study of food variety seeking in childhood, adolescence and early adult life. *Appetite,* 44: 289-297.

692 Pagliarini, Ella, Nicola Gabbiadini, and Sabrina Ratti. 2005. Consumer testing with children on
693 food combinations for school lunch. *Food Quality and Preference,* 16: 131-138.
694 Pliner, Patricia. 1994. Development of measures of food neophobia in children. *Appetite,* 23:
695 147-163.
696 Pliner, Patricia, and Karen Hobden. 1992. Development of a scale to measure the trait of food
697 neophobia in humans. *Appetite,* 19: 105-120.
698 Pliner, Patricia, and Sarah J. Salvy. 2006. *Food neophobia in humans*. In R. Shepherd, M. Raats
699 (Eds.) The psychology of food choice (75-92). Wallingford, Oxfordshire: CABI
700 Publishing.
701 Raju, S., Rajagopal, P., and Gilbride, T. J. 2010. Marketing healthful eating to children: the
702 effectiveness of incentives, pledges, and competitions. *Journal of Marketing*, 74(3): 93-
703 106.
704 Robertson, Ruth 2008. *Using information to promote healthy behaviours*, King's Fund Report.
705 Russell, Catherine G., and Anthony Worsley. 2008. A population-based study of preschoolers'
706 food neophobia and its associations with food preferences. *Journal of Nutrition*
707 *Education and Behavior,* 40: 11-19.
708 Schofield, Heather, George Loewenstein, Jessica Kopsic, and Kevin G. Volpp. 2015. Comparing
709 the effectiveness of individualistic, altruistic, and competitive incentives in motivating
710 completion of mental exercises. *Journal of Health Economics,* 44: 286-299.
711 Smith, Trenton, and Attila Tasnadi. 2007. A theory of natural addiction. *Games and Economic*
712 *Behavior,* 59: 316–344.
713 van Kleef, E., Vrijhof, M., Polet, I.A., Vingerhoeds, M. H., de Wijk, R. 2014. Nudging children
714 towards whole wheat bread: a field experiment on the influence of fun bread roll shape
715 on breakfast consumption. *BMC Public Health,* 14(906): 1-11.

716 van Kleef, E. Bruggers, I., de Vet, E. 2015. Encouraging vegetable intake as a snack among
717 children: the influence of portion and unit size. *Public Health Nutrition*, 18(15): 2736-
718 2741.
719 Volpp, Kevin G., Leslie John, Andrea B. Troxel, Laurie Norton, Jennifer Fassbender, and George
720 Lowenstein. 2008. Financial incentive–based approaches for weight loss: a randomized
721 trial. *Journal of the American Medical Association,* 300: 2631-2637.
722 Volpp, Kevin G., Andrea B. Troxel, Mark V. Pauly, Henry A. Glick, Andrea Puig, David A. Asch,
723 Robert Galvin, Jingsan Zhu, Fei Wan, Jill DeGuzman, Elizabeth Corbett, Janet Weiner,
724 and Janet Audrain-McGovern. 2009. A randomized, controlled trial of financial
725 incentives for smoking cessation. *The New England Journal of Medicine,* 360: 699-709.
726 Yu, Hyunjae. 2011. Parental Communication Style's Impact on Children's Attitudes Toward
727 Obesity and Food Advertising. *Journal of Consumer Affairs*, 45(1): 87-107.
728 World Health Organization. 2009. *Global Health Risks: Mortality and Burden of Disease*
729 *Attributable to Selected Major Risks*. World Health Organization, Geneva.
730

TABLE 1

Incentive liking scores

Prizes **1st -5th grade**	**Liking scores**	**Prizes** **6th -8th grade**	**Mean Liking scores**
Emoji keychains	28	Headphones	4.42
Sticky hands	25	Usb cable	4.1
Pocket ball game	20	Backpack	3.73
Water pistol	16	Smartphone holder	3.68
Monster hand-puppets	13	Pocket ball game	2.71
Fluorescent skeleton	12	Photobook stickers	2.65
Fluorescent bugs	12	Funny sunglasses	2.19

TABLE 2

T-test for children's vegetable liking and food neophobia across the two groups

	Group	Mean	Obs	p-value
Vegetable liking	Control	3.74	157	0.147
	Incentive	3.64	156	
Salad	Control	4.28	157	0.931
	Incentive	4.38	156	
Green beans	Control	4.59	32	0.339
	Incentive	4.52	32	
Broccoli	Control	3.57	32	0.156
	Incentive	3.42	32	
Vegetable mix	Control	3.11	32	0.803
	Incentive	3.22	31	
Carrots	Control	3.94	29	0.050
	Incentive	3.61	29	
Spinach	Control	3.51	32	0.299
	Incentive	3.4	32	
Neophobia	Control	2.56	157	0.774
	Incentive	2.58	156	

TABLE 3

Food neophobia scale items

Item	Description	Obs.	Mean	SD	Min	Max
1	I eat almost every day new and unusual foods (R)	371	2.90	1.14	1	5
2	I don't trust new foods	370	2.52	1.20	1	5
3	If a food is new, I don't try it	369	2.27	1.23	1	5
4	I like to try weird tastes and foods, which are unusual and coming from different countries (R)	369	2.60	1.34	1	5
5	When I am at a friend's party, I like to try new food (R)	367	2.16	1.15	1	5
6	I am afraid to eat food I have never had before	368	2.41	1.24	1	5
7	I am very fussy when it's a matter of food	366	2.44	1.28	1	5
8	I really eat everything! (R)	367	3.17	1.29	1	5

TABLE 4

T-test for whether vegetable leftovers differ across the control and incentive groups at T_0, T_1, T_2, and T_3 respectively

Period	**Group**	**Leftover (g/day)**		
		Mean	Obs	p-value
T_0	Control	659.9	148	0.200
	Incentive	633.3	151	
T_1	Control	727.9	157	0.001
	Incentive	473.6	156	
T_2	Control	607.0	102	0.008
	Incentive	516.1	104	
T_3	Control	554.4	40	0.004
	Incentive	430.6	40	

TABLE 5

Regression results for the effect of the treatment on vegetable leftovers, respectively controlling for age, vegetable variety and food neophobia.

	Model 1	**Model2**	**Model 3**	**Model 4**	**Model 5**
	Leftover (g/day)	Leftover (g/day)	Leftover (g/day)	Leftover (g/day)	Leftover (g/day)
Treatment	-0.320 ***	-0.086 *	-0.087 **	-0.098 **	-0.113 **
	(0.043)	*(0.047)*	*(0.040)*	*(0.038)*	*(0.039)*
Treatment*T_1		-0.465 ***	-0.462 ***	-0.453 ***	-0.454 ***
		(0.096)	*(0.080)*	*(0.079)*	*(0.078)*
Treatment*T_2		-0.157 **	-0.158 **	-0.130 **	-0.131 **
		(0.070)	*(0.058)*	*(0.057)*	*(0.059)*
Treatment*T_3		-0.422 ***	-0.423 ***	-0.414 ***	-0.414 ***
		(0.112)	*(0.097)*	*(0.010)*	*(0.102)*
Grade 1 [a]			-	-	-
Grade 2			0.336 ***	0.341 ***	0.453 ***
			(0.051)	*(0.049)*	*(0.056)*
Grade 3			-0.013	-0.011	0.187
			(0.051)	*(0.048)*	*(0.066)*
Grade 4			-0.200 ***	-0.196 ***	-0.168 ***
			(0.049)	*(0.046)*	*(0.046)*
Grade 5			-0.735 ***	-0.732 ***	-0.421 ***
			(0.102)	*(0.100)*	*(0.085)*
Grade 6			0.282 ***	0.286 ***	0.257 ***
			(0.044)	*(0.042)*	*(0.043)*
Grade 7			-0.068	-0.068	0.074
			(0.060)	*(0.055)*	*(0.070)*
Grade 8			-0.061	-0.058	-0.031
			(0.074)	*(0.073)*	*(0.078)*
Green bean [a]				-	-
Broccoli				0.032	0.031
				(0.057)	*(0.055)*
Veg. mix				-0.189 ***	-0.190 ***
				(0.048)	*(0.047)*
Carrots				-0.195 ***	0.195 ***
				(0.049)	*(0.049)*
Spinach				0.154 ***	0.153 ***
				(0.045)	*(0.045)*
Neophobia					0.566 **
					(0.138) ***
Constant	6.422 ***	6.422 ***	6.481 ***	6.520 ***	4.973 ***
	(0.020)	*(0.020)*	*(0.040)*	*(0.048)*	*(0.053)*
Observations	898[b]	898	898	898	898
F	55.46	17.54	29.06	28.16	27.07
Prob > F	0.000	0.000	0.000	0.000	0.000
R-squared	0.058	0.105	0.320	0.360	0.377
Root MSE	0.647	0.631	0.553	0.537	0.530

Notes:

[a] Removed for estimation purposes; Robust standard error in parentheses, significance at p<0.05*, p<0.01**,p<0.001***

[b] The total number of observation is given by daily data at class level over the 12 experimental weeks excluding festive days

FIGURE 1

Graphic representation of vegetable leftover patterns over the 11 experimental weeks and the follow up.

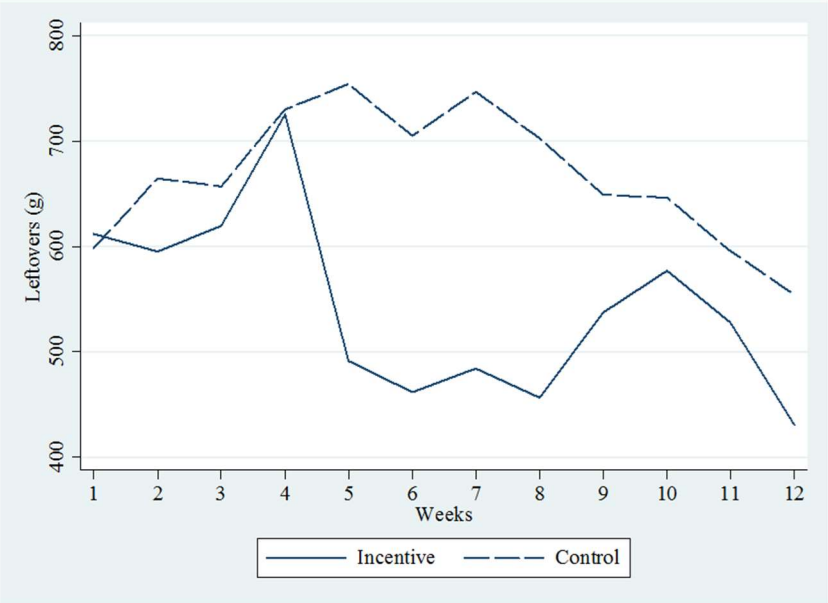
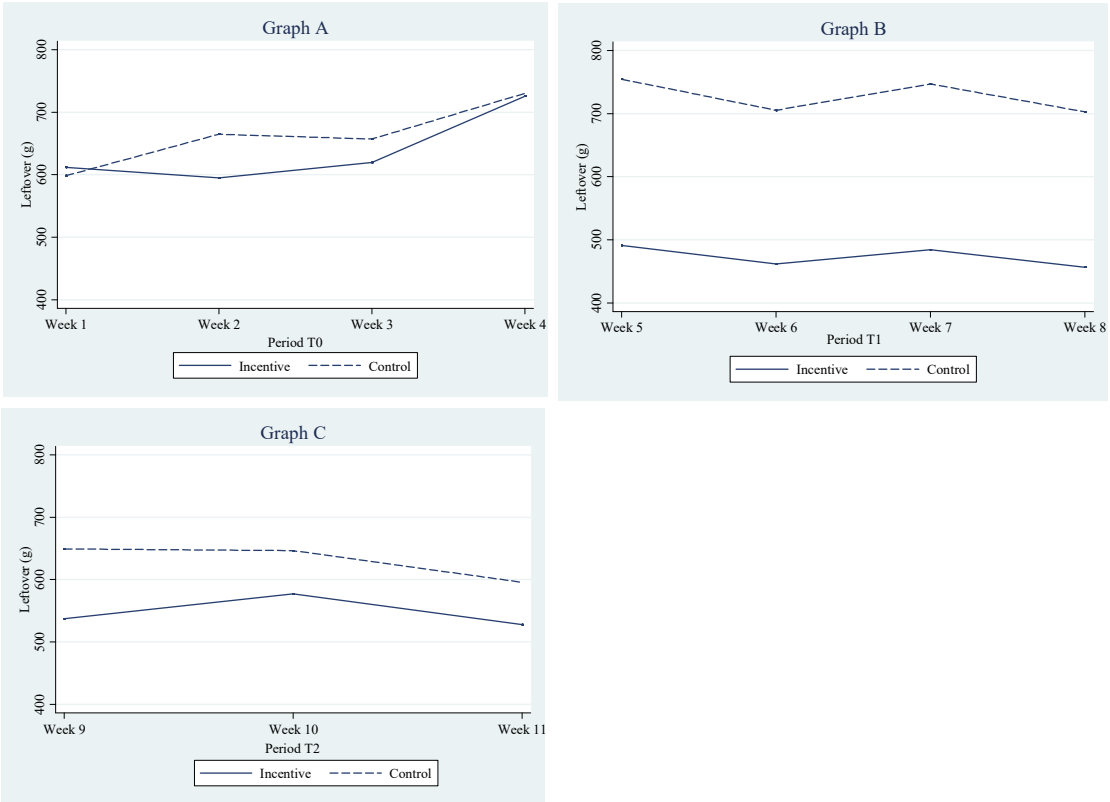


FIGURE 2

Graphic representation of vegetable leftovers of the control and incentive groups, respectively at T_0, T_1, and T_2



APPENDIX A – EXPERIMENTAL MATERIALS

Non-monetary incentives

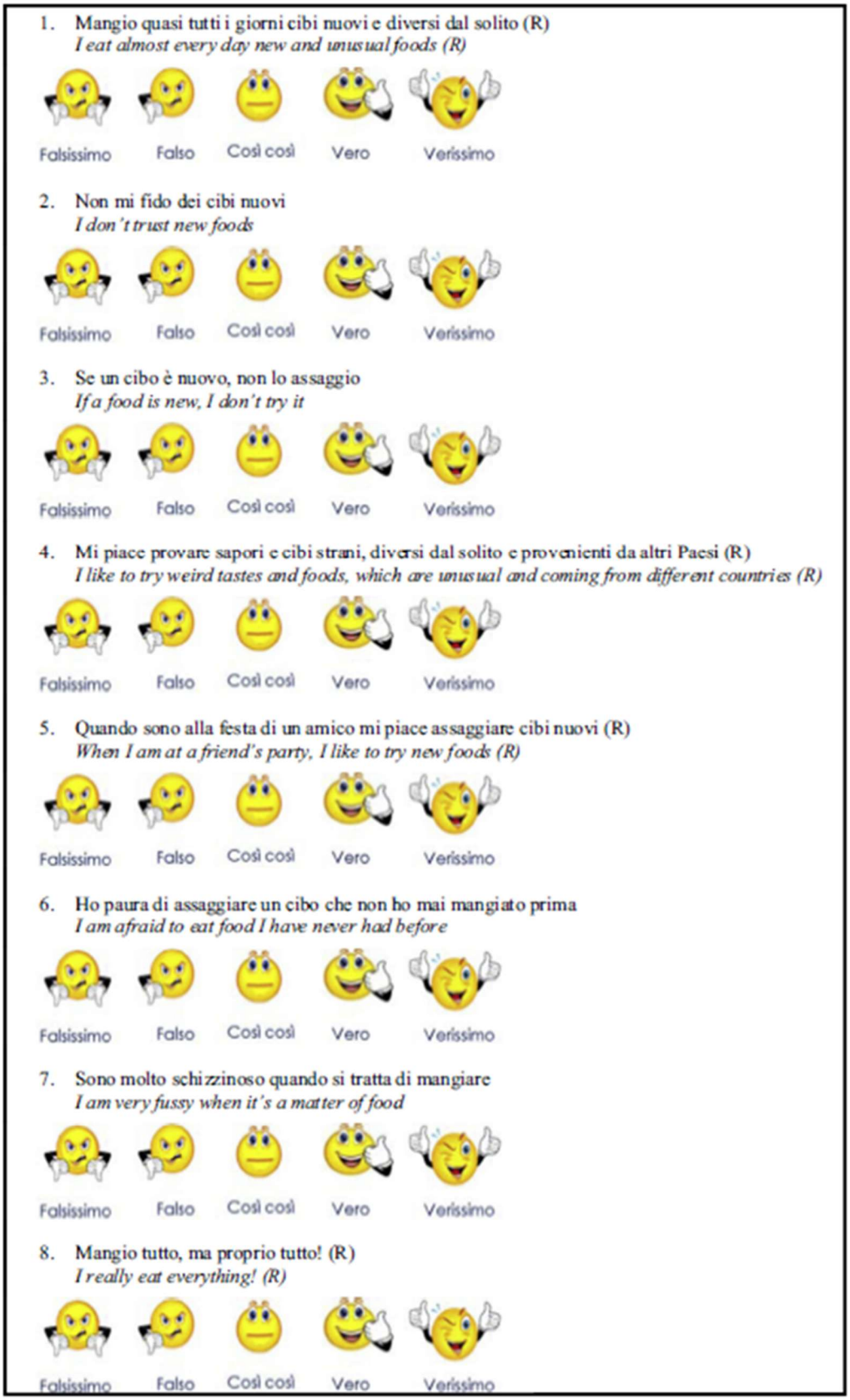
Non-monetary incentives - 1st to 5th grade	
	Emoji keychains
	Sticky hands
	Pocket ball game
	Water pistol

Non-monetary incentives - 6th to 8th grade	
	Headphones

	Usb cable
	Backpack
	Smartphone holder

Italian Child food Neophobia Scale (ICFNS)

(Laureati, M., Bergamaschi, V., Pagliarini, E., 2015. Assessing childhood food neophobia: Validation of a scale in Italian primary school children. Food Quality and Preference 40, 8-15).



1. Mangio quasi tutti i giorni cibi nuovi e diversi dal solito (R)
   I eat almost every day new and unusual foods (R)

   Falsissimo | Falso | Così così | Vero | Verissimo

2. Non mi fido dei cibi nuovi
   I don't trust new foods

   Falsissimo | Falso | Così così | Vero | Verissimo

3. Se un cibo è nuovo, non lo assaggio
   If a food is new, I don't try it

   Falsissimo | Falso | Così così | Vero | Verissimo

4. Mi piace provare sapori e cibi strani, diversi dal solito e provenienti da altri Paesi (R)
   I like to try weird tastes and foods, which are unusual and coming from different countries (R)

   Falsissimo | Falso | Così così | Vero | Verissimo

5. Quando sono alla festa di un amico mi piace assaggiare cibi nuovi (R)
   When I am at a friend's party, I like to try new foods (R)

   Falsissimo | Falso | Così così | Vero | Verissimo

6. Ho paura di assaggiare un cibo che non ho mai mangiato prima
   I am afraid to eat food I have never had before

   Falsissimo | Falso | Così così | Vero | Verissimo

7. Sono molto schizzinoso quando si tratta di mangiare
   I am very fussy when it's a matter of food

   Falsissimo | Falso | Così così | Vero | Verissimo

8. Mangio tutto, ma proprio tutto! (R)
   I really eat everything! (R)

   Falsissimo | Falso | Così così | Vero | Verissimo

4. Vegetable liking questionnaire

Class: _____

(Mark the face that corresponds to your answer)

How much you like these vegetables?

Salad

Super good (Mitic)	Really Good (I like it very much)	Good (I like it)	So so (Neither good, nor bad)	Bad (I don't like it)	Really bad (I really don't like it)	Super bad (You eat it)

Broccoli

Super good (Mitic)	Really Good (I like it very much)	Good (I like it)	So so (Neither good, nor bad)	Bad (I don't like it)	Really bad (I really don't like it)	Super bad (You eat it)

Vegetable mix

Super good (Mitic)	Really Good (I like it very much)	Good (I like it)	So so (Neither good, nor bad)	Bad (I don't like it)	Really bad (I really don't like it)	Super bad (You eat it)

Carrots

Super good (Mitic)	Really Good (I like it very much)	Good (I like it)	So so (Neither good, nor bad)	Bad (I don't like it)	Really bad (I really don't like it)	Super bad (You eat it)

Green Beans

Super good (Mitic)	Really Good (I like it very much)	Good (I like it)	So so (Neither good, nor bad)	Bad (I don't like it)	Really bad (I really don't like it)	Super bad (You eat it)

895  **Spinach**

Super good (Mitic)	*Really Good (I like it very much)*	*Good (I like it)*	*So so (Neither good, nor bad)*	*Bad (I don't like it)*	*Really bad (I really don't like it)*	*Super bad (You eat it)*