Incentivizing Nutritious Diets:
A Field Experiment of Relative Price Changes and How They are Framed[1]

John Cawley, Cornell University
Andrew Hanks, Ohio State University
David Just, Cornell University
Brian Wansink, Cornell University

July 18, 2015

Abstract

This paper seeks to answer the following research questions: 1) Are consumers' food purchases responsive to less-nutritious food being made relatively more expensive than nutritious food? 2) Does consumer responsiveness depend on whether the price change is framed as a tax on less-nutritious food, a subsidy for nutritious food, or both? 3) Do the answers differ by the income or education of the consumer? We answer these questions using a randomized controlled field experiment that involved 208 households and which lasted eight months. Nutritious food (classified according to an existing supermarket shelf-label nutrition guidance system) was made 10% cheaper than less-nutritious food, with food purchases tracked using supermarket scanner cards and subsidies paid electronically via debit cards. The results indicate that, overall, the 10% relative price difference did not significantly affect purchases. However, we find significant differences by socioeconomic status; low-income households respond to the subsidy frame by buying more of both nutritious and less-nutritious food. This implies that attempts to subsidize nutritious food for low-income households may lead them to also buy more of what public health wishes to discourage: less-nutritious food.

[1] We gratefully acknowledge financial support from the National Institutes of Health (NIH) grant 1RC1HD063370-01. The NIH played no other role in the conduct of the study. The Cornell University Institutional Review Board approved the design of this study (Protocol ID#1110002491). For helpful comments and suggestions, we thank Heather Royer and participants at the American Society of Health Economists biennial conference, the International Health Economics Association conference, the TIGER conference in Toulouse France, and seminar participants at the Indiana University, McGill University, University of Oxford, the University of Pennsylvania, and the University of Sydney. Contact author: John Cawley, 2312 MVR Hall, Cornell University Ithaca NY 14853. JHC38@cornell.edu

Introduction

Poor diets in the United States and many other economically developed countries contribute to high rates of chronic disease. For example, in the U.S., 37% of the adult population has cardiovascular disease, 16% has high total blood cholesterol, 34% has hypertension, 11% has diabetes, and it is estimated that 41% will be diagnosed with some form of cancer during their lifetime (USDA, 2010). Moreover, 35.1% of adults and 16.9% of youths in the U.S. are obese (Ogden et al., 2014).

Diet-related chronic disease is a global problem. Worldwide, the annual deaths due to high blood pressure total 7.5 million, high blood glucose (diabetes) 3.4 million, overweight and obesity 2.8 million, and high cholesterol 2.6 million (WHO, 2009). Even in low-income countries, the top 10 risk factors for preventable death include high blood pressure, high blood glucose, and high cholesterol (WHO, 2009).

The problems with many modern diets, which contributes to these high rates of chronic disease, are that they contain too much saturated fats, trans fats[2], cholesterol, added sugars, added sodium, and refined grains, and too little whole grains and fresh fruits and vegetables (USDA, 2010).

As a result of the high rates of chronic disease, there have been calls for taxes on energy-dense less-nutritious foods from many medical and public health organizations, such as the World Health Organization (2015), U.S. Dietary Guidelines Advisory Committee (2015), British Medical Association (2015), Institute of Medicine (2009), and the International Obesity Task Force (2005), which urged all European Union member countries to enact taxes on energy-dense

[2] Trans fat has become far less common in the U.S. diet. Since 2006, when the U.S. Food and Drug Administration required that trans fat content be included in the Nutrition Facts panel, food manufacturers have reformulated to remove trans fats from their products (Mozaffarian et al., 2010) and levels of trans fats in the blood have fallen 58% (Vesper et al., 2012). In June, 2015, the FDA announced that food manufacturers had three years to phase out the use of trans fats in food.

foods. There have also been numerous calls in medical journals for taxes to incentivize a healthy diet (e.g. Brownell and Frieden, 2009, and Jacobson and Brownell, 2000). Taxes on energy-dense foods are arguably the most commonly-advocated anti-obesity policy.

Policymakers have responded to this call for action. Numerous countries, such as Australia, Canada, Denmark, Fiji, Finland, France, Hungary, Norway, and Mexico, have recently implemented taxes on energy-dense, less-nutritious foods (see e.g. World Health Organization, 2015, Sassi et al., 2013, and Thow et al., 2011). In the U.S., 34 states tax soft drinks sold in grocery stores, at an average rate of 4.02%, and 15 states tax snacks sold in grocery stores at an average rate of 1.2% (Chriqui et al., 2008). In early 2015, Berkeley, California became the first U.S. city to impose an excise tax on sugar-sweetened beverages.

To some extent, an individual's diet and any resulting chronic disease or premature mortality can be seen as a private, individual decision. However, there are two economic rationales for government intervention to incentivize healthier diets. First, there are external costs of a poor diet that operate through private and public health insurance (Cawley, 2015). Premiums that fund private health insurance, and the taxes that fund public health insurance, are not a function of diet, and as a result, the costs of treating diet-related chronic disease are borne not only by those with the disease but also by others in the same insurance pools and by taxpayers. The exact magnitude of these external costs is not known, but they are undoubtedly large given the enormous medical care costs. It is estimated that the annual direct medical care costs total $273 billion for cardiovascular disease (CDC, 2015a), $315.8 billion for obesity (Cawley et al., 2015), $116 billion for diabetes (CDC, 2015b), and $263.8 billion for cancer (this includes both direct and indirect costs; CDC 2015c). Clearly, to pool these separate estimates would involve some degree of over-counting, but the overall cost of these diseases is clearly very high.

Behavioral economics offers a second rationale for government intervention to incentivize healthier diets. Individuals may have time-inconsistent preferences; they may want to eat a nutritious diet so as to be healthy in the future, but in the short run may be tempted by the immediate gratification (Laibson, 2014). Some have argued that optimal taxes should reflect not only externalities but also internalities associated with time-inconsistent preferences, and that in such cases sin taxes can make those who engage in such activities happier because it helps them help themselves (Gruber and Mullainathan, 2005).

It is difficult to estimate the effect of existing food taxes on purchases and consumption. In the U.S., state-level taxes are so small that it is very difficult to measure their effects (Fletcher, Frisvold, and Tefft, 2010; Chaloupka et al., 2011; Fletcher et al., 2011). For national taxes, it is difficult to disentangle the effect of the tax from time effects. In other words, it is hard to identify a geographic control group. For both, policy endogeneity is a problem.

A series of field experiments have estimated consumer responsiveness to price changes, but have often paired those price changes with related interventions such as signs or marketing, the effect of which is confounded with the price change. For example, a set of experiments conducted by researchers at the University of Minnesota manipulated prices in cafeterias and vending machines and found that a 50% subsidy for fruits and salads tripled sales, but sales fell to baseline after the subsidy was removed (French et al., 1997; Jeffrey et al., 1994). Elbel et al. (2013) opened their own store in a hospital, and imposed a 30% tax on unhealthy foods, which they juxtaposed next to healthier alternatives. They estimate that the tax increased the probability of consumers choosing healthier alternatives by 11 percentage points. The generalizability is questionable given that the store was a researcher-created environment that involved deliberate juxtapositioning of healthier and less healthy options. The USDA conducted the Healthy Incentives Pilot for recipients of the Supplemental Nutrition Assistance Program (SNAP), which offered a 30 cent rebate to the Electronic Benefit Transfer card for each dollar spent on fruits and vegetables. It estimated that the program resulted in 0.22 cups/day more fruits and vegetable consumed by participating adults (USDA, 2013).

The contribution of this paper is to estimate the responsiveness of consumers to a price change – with no other interventions such as additional signage or juxtapositioning of alternatives – in the consumer's usual retail environment. In other words, we observe consumers buying their usual items in the supermarket in which they typically shop. Specifically, we conduct a randomized controlled field experiment in order to measure the impact of a 10% relative price difference between nutritious and less-nutritious food in order to answer three research questions: 1) Are consumers' food purchases responsive to less-nutritious food being made 10% more expensive than nutritious food? 2) Does that responsiveness depend on whether the price change is framed as a tax on less-nutritious food, a subsidy for nutritious food, or both? 3) Do the answers differ by the education or income of the consumer?

We hypothesize that the framing of the relative price change as a tax or subsidy may affect consumer response in light of prospect theory, which states that people interpret gains and losses relative to a reference point (Kahneman and Tversky, 1979). In particular, people may respond more when the tradeoff is framed as a loss rather than a foregone reward (Gachter et al., 2009; Homonoff, 2015), which suggests that people may be more responsive to the frame of a tax on less-nutritious food than that of a subsidy for nutritious food.

Additionally, we hypothesize that responses to the relative price change may differ by socioeconomic status. First, consumer response may differ by income for several reasons. Mullainathan and Shafir (2013) argue that poverty consumes mental bandwidth, which implies that lower-income individuals may pay less attention to the price change. On the other hand, other evidence suggests that lower-income individuals may be more responsive to the relative price change. Low-income individuals who receive public assistance (such as food stamps or social security) exhibit "first of the month effects" – their spending on food decreases as the month progresses (Hastings and Washington, 2010; Shapiro, 2005). This suggests that they may be credit constrained and perhaps price reductions could have substantial income effects. Other research suggests that the income elasticity of body weight is greater for low-income individuals (Akee et al., 2013; Schmeiser, 2009).

Second, consumer response may also differ by education. The better educated tend to demand more health and be more efficient producers of their own health (Grossman, 1972) and thus may have a more elastic demand for nutritious food. In addition, the better educated may simply better understand the treatment or respond to changing prices in general.

This paper builds on previous studies that conducted field experiments manipulating prices. A review of the literature by Epstein et al. (2012) finds only four studies that manipulated prices of foods in supermarkets; all provided discounts for healthy foods (none increased the relative price of unhealthy foods), and three of the four examined only purchases of a subset of available foods and thus there was little evidence of substitution patterns. Other experiments manipulating food prices took place in laboratories, cafeterias and restaurants, farmer's markets, and vending machines (Epstein et al., 2012). In many cases, the price treatment was combined with other treatments, such as signs alerting consumers to the price change and juxtaposing healthy with unhealthy products to facilitate comparisons. The contributions of the present study to the existing literature are: it takes place in an existing supermarket, no other treatment than the price change takes place (i.e. no rearranging of items, no additional signs), we observe all food purchases made at the supermarket (and provide incentives for subjects to do all of their food shopping at the supermarket), and we rely on an objective system that classifies food as nutritious and less-nutritious and which is already in place in the supermarket.

Data and Methods

The Field Experiment

Controlled field studies with random assignment have the potential to clearly identify causal effects (List, 2009). Between May 1 and June 30, 2010, we recruited 239 loyalty card shoppers to participate in the study. Individuals were recruited via face-to-face contact at the entrances to two grocery stores in upstate New York. These stores are part of a regional supermarket chain that is located in the Northeast U.S. In order to ensure a diverse set of participants, subjects were recruited at various days and times, as well as at two different stores of the same chain in neighborhoods of differing socioeconomic status. In addition, to be eligible for inclusion in the study, participants had to have children under the age of 18 years living at home, do at least 75% of their shopping at the supermarket chain, and do a majority of the household's shopping.

After enrollment, subjects were sent an email with a link to complete a survey on their household characteristics and shopping patterns. After repeated requests, fourteen subjects did not complete the survey and were dropped. One household later attrited from the study and so we drop all data from that household. Furthermore, in 16 households, two individuals claimed to each do half of the household's shopping. Both were enrolled but purchases were aggregated to the household level. As a result, we have complete information, survey responses and expenditure data, for 208 households.

Participating households received two cards. A scanner card (with the subject's name and photograph) was used to track purchases at the supermarket checkout lane. A debit card was used to deliver incentives and subsidies, which were electronically credited on a weekly basis.

Any experiment designed to manipulate the prices of nutritious and less-nutritious foods faces the challenge of defining those two categories. We relied upon a supermarket shelf-label nutrition guidance system that had already been in place in the supermarket for several years prior to this experiment. This proprietary system, called Guiding Stars, scores foods based on their nutritional value. More specifically, it takes into account vitamins, minerals, fiber and whole grains (which raise the score) and saturated fat, trans fat, cholesterol and added sugar and sodium (which lower the score). Ultimately, foods are rated on a scale from zero stars (poor nutritional value) to three stars (best nutritional value), and this score is displayed on the supermarket shelf label below each food item (retail price and unit price). Over 60,000 food items are rated. The few foods that are not rated are new (and thus not yet rated), seasonal (not consistently available), or have no calorie or nutrient content (such as dried spices or dried coffee or tea). For more information on Guiding Stars, see Fischer et al. (2011).

For our experiment, we defined less-nutritious food as that which receives zero stars, and nutritious food as that which receives any stars (one, two, or three). An incentive scheme could offer more finely-tuned subsidies based on whether the item received one, two, or three stars, but that would also involve the tradeoff of increased complexity that could cause confusion for study participants. We chose to make the intervention simple to understand, and divided foods into those with zero stars (which were made relatively more expensive) and those with any stars (which were made relatively cheaper).[3] For simplicity, we refer to these groups throughout this paper as less-nutritious foods and nutritious foods.

We observed households' food purchases during an eight-week baseline period before altering the relative prices of nutritious and less-nutritious foods.[4] To encourage households to conduct all of their food shopping at the participating supermarket, they received a 10% discount on purchases of all rated food items, that is, on any foods rated with 0, 1, 2, or 3 stars.

At the conclusion of the baseline period, subjects were randomized into one of four groups. The control group (N=52 households) continued to receive a 10% discount on all rated food items. For the treatment group (N=156), nutritious food was made 10% cheaper than less-nutritious food. How this price wedge was framed differed based on the treatment group into which the subject was randomized. The tax group (N=51) was told that they received a 15% discount on all rated food items, but were taxed 10% (and thus received only a 5% discount) on less-nutritious food. The subsidy group (N=55) was told it received a 5% discount on all rated food items, plus an additional 10% subsidy on nutritious food, for a total of 15% off nutritious food. The tax/subsidy group (N=50) was told that it received a 10% discount on all rated food items, plus an additional 5% subsidy on nutritious food (for a total subsidy of 15%) but was taxed 5% on less-nutritious food (for a net subsidy of 5%). In all three treatment conditions, nutritious food was subsidized 15% and less-nutritious food was subsidized 5%; thus each group

[3] The prices of unrated items were not altered.

[4] Households signed up 5-8 weeks before the treatment period; we dropped the first 3 weeks during which some but not all households were enrolled, and use as the baseline period the final 4 weeks before the treatment period when all households were enrolled.

faced a 10% price wedge between nutritious and less-nutritious food, and the only way the treatment differed was how it was framed.

Households were notified of their respective treatment via email and phone calls. Out of concern that subjects may not check their email or voice messages, the enrolled representative from each household was also individually contacted by phone and spoken to directly, a process that took 12 days. We removed these two weeks from analysis because some subjects may not have yet been aware of their treatment condition.

Note that in a voluntary field experiment experimenters cannot impose taxes on less-nutritious foods greater than the participation incentive or subjects would likely shop elsewhere to buy these foods, and such expenditures would not be recorded by the study. To address this, the participation incentive was always greater than the tax imposed, ensuring that shoppers could not be worse off by shopping at the study stores. Because the participation incentive was also offered during the baseline period, we were able to identify the effect of price changes from the relative price changes between nutritious and less-nutritious foods that were imposed between the baseline and treatment periods. See Table 1 for the relative price changes at baseline and during the treatment period, and details of the framing of the treatment.

To clarify, prices on the supermarket shelves were not altered. The participating supermarket was understandably unwilling to allow the researchers to manipulate shelf prices for all of their customers. Instead, subjects' purchases were tracked using the scanner cards, and the discounts, net of taxes, were uploaded weekly to the debit card. Each subject received a weekly email notifying them of the amount of incentive or subsidy they had received, and reminding them which foods were taxed and which were subsidized.

The treatment period lasted for 25 weeks and ended without prior notice. See Figure 1 for a detailed timeline of the study.

Data

Itemized grocery purchases of each subject were tracked by the supermarket using the scanner cards. The item-level transaction data include: date, quantity of item, expenditures on item, Guiding Stars score of each item, and the description of the product. These transactions were aggregated to the level of household and week, with weeks defined as Monday through Sunday. We merge the information from the baseline survey to the transaction data.

We focus on two main outcomes: the household's expenditures (defined before any subsidies or taxes applied by the experiment) and quantity purchased. Quantity purchased is measured in units, which is a limited measure because it does not account for size differences. For example, a half gallon and a gallon of milk each count as one unit, as do two different-sized boxes of the same cereal. Thus, this measure of quantity is a noisy measure of the quantity of food purchased. We examine these two outcomes for all food purchases, as well as separately for nutritious food and less-nutritious food.

If a household did not buy any food in that category in that week, the values of expenditures and quantity purchased are set to zero. The exception to this occurred during the first three weeks of the baseline period when households were still matriculating in the study. During these three weeks, instances with no expenditures were treated as missing until the household had a recorded shopping trip.

Hypotheses and Empirical Methods

We test the following hypotheses:

H1: Increasing the price of less-nutritious food relative to the price of nutritious food will decrease purchases of less-nutritious food and increase purchases of nutritious food;

H2: Framing the relative price change as a subsidy for nutritious food will increase the extent to which the price change increases purchases of nutritious food, and framing the relative price change as a tax on less-nutritious food will increase the extent to which the relative price change decreases purchases of less-nutritious food;

H3: These effects will vary by income and education.

In order to test these hypotheses, we estimate fixed effects difference-in-differences models of expenditures and quantities. Randomization into the treatment and control groups allows for interpretation of the difference-in-differences estimator as a causal effect of the treatment.

To estimate the average effect of the price change, ignoring the possibility of framing effects, we estimate the following regression model:

$$y_{hw} = \alpha_0 + \beta_0 Treatment_h * Post_w + \sum_{w=1}^{W} \chi_0 \mathrm{I}_w + \sum_{h=1}^{H} \delta_0 \mathrm{I}_h + \varepsilon_{hw} \qquad 1$$

The data are at the level of household (*h*) and week (*w*). The difference-in-differences estimator is β_0, the coefficient on the interaction term of being in the treatment group and the week in question being after the treatment began. This coefficient measures the change between the baseline and treatment period for the treatment group relative to the control group. In order to control for time-invariant unobserved heterogeneity among households, the model controls for a full set of household fixed effects I_h . In order to control for time effects (such as the seasonal availability of fresh fruits and vegetables and changes in demand due to holidays), the model controls for a full set of week fixed effects I_w. The OLS regression model is estimated for all food purchases, as well as separately for purchases of nutritious food and less-nutritious food. The null hypothesis is that the 10% price wedge has no impact on purchases: $\beta_0=0$. To account for possible correlation in errors for the same household over time, standard errors are clustered by household.

In order to test whether the framing of the price change affects consumers' response to the price change, we estimate the following model, which estimates a separate difference-in-differences effect for each of the three treatment groups (tax, subsidy, tax and subsidy):

$$\begin{aligned} y_{hw} = \alpha_1 + \beta_1 Tax_h * Post_w + \beta_2 Subsidy_h * Post_w + \beta_3 Tax \,\&\, Subsidy_h * Post_w \\ + \sum_{w=1}^{W} \chi_1 \mathrm{I}_w + \sum_{h=1}^{H} \delta_1 \mathrm{I}_h + \varepsilon_{hw} \end{aligned} \qquad 2$$

The null hypothesis is that the framing of the treatment as either a tax on less-nutritious food, a subsidy of nutritious food, or both, does not alter the treatment effect; i.e. that $\beta_1= \beta_2= \beta_3$.

To test whether the treatment effect varies by income, we estimate models 1 and 2 separately for those whose household income is a) below or b) above 130% of the Federal Poverty Line (FPL), which is the eligibility threshold for the Supplemental Nutrition Assistance Program (SNAP) and is close to the eligibility threshold for Medicaid (133% of FPL).

To test whether the treatment effect varies by education, we estimate the model separately for those whose educational attainment is a) a high school degree or less or b) some college or more.

We emphasize that, given our overall sample size, we have limited statistical power for subgroups. When we divide the sample by income, we have 36 households below and 155 households above, 130% of the FPL. When we divide the sample by education, we have 18 participants with a high school education or less, and 182 participants with some college or more education (see Table 2). These subtotals do not sum to our total of 208 households because of non-response to the questions about income and education.

Empirical Results

Summary Statistics

Tables 2 and 3 list summary statistics for the study participants, with columns for the whole sample, control group, all treatment groups pooled, and each treatment group separately. Table 2 reports sample sizes for the socioeconomic subgroups. Table 3 reports summary statistics for additional household characteristics, such as income, number of children at home, household size, marital status, and race/ethnicity, which were all controlled for in the fixed effects regression equations.

The summary statistics indicate that our sample is relatively well educated (91% have more than a high school education) and high income (19.0% have an income over $100,000), and is 93.7% white. This is a reflection of the fact that the participating supermarket chain is relatively high-quality and that our sample consists of individuals in upstate New York. By construction, the sample consists of families with at least one child under the age of 18 years in the household.

Table 4 lists unconditional weekly expenditures on foods (overall, all rated, less nutritious, nutritious) for the entire sample and by group (control, all treatment, each treatment group). Household weekly food expenditures at this supermarket averaged $89.83 during the baseline period, and $100.88 during the treatment period. In comparison, data from the Consumer Expenditure Survey indicate that on average U.S. households spent $76 per week on food purchased for at-home consumption in 2013 (BLS, 2015). Notably the BLS estimate is unconditional, whereas our sample consists of households with at least one child under the age of 18 years, and are thus likely to be above-average in terms of food expenditures.

The increase in average weekly food expenditures for all treatment groups ($10.95) is roughly equal to that for the control group ($11.32); this unconditional difference-in-differences suggests that the treatment did not significantly affect overall expenditures on food. The increase in expenditures on nutritious food specifically was also similar for all treatment groups pooled ($4.69) and the control group ($3.30).

Unconditional weekly quantities purchased, and the shares of the purchases that were of nutritious food, are listed in Appendix Tables 1 and 2.

Overall Effect of Relative Price Change

Table 5 lists results of the difference-in-differences models for expenditures and quantities. Our hypothesis is that the 10% relative price change increased the quantity demanded of nutritious food, and decreased the quantity demanded of less-nutritious food. Table 5 shows that the point estimates of the coefficients are consistent with that those hypotheses, but the coefficients are not statistically significant. For example, we find that creating a 10% price difference between nutritious and less-nutritious foods raised spending on nutritious food by $1.11 per week and lowered spending on less nutritious food by $1.55 per week, neither of which was statistically significant. On net, spending on food rated by Guiding Stars fell by $0.44 per week, which was not statistically significant. In terms of quantities, the 10% relative price difference increased purchases of nutritious food by 0.95 units and lowered purchases of less nutritious food by 0.87 units; overall purchases of foods rated by Guiding Stars rose by .08. None of those changes are statistically significant.

In summary, we are unable to reject the null hypothesis of no effect of the relative price change on purchases of nutritious and less-nutritious foods.

Effect of Framing of Relative Price Change

Next we test whether the effect of the relative price change differed by the way in which it was framed: as a tax on less-nutritious food, a subsidy for nutritious food, or both. It is possible that, because of loss aversion, the tax frame may exhibit greater treatment effect than the subsidy frame. Moreover, given the difference in salience, we may see a greater increase in purchases of nutritious food for the subsidy frame, but a greater decrease in purchases of less-nutritious food for the tax frame.

Table 6 presents the results of the difference-in-difference models that estimate separate effects by frame. In no case are the treatment effects significantly different across frames (whether tax versus subsidy, tax versus tax/subsidy, or subsidy versus tax/subsidy). In addition, no estimated treatment effect for nutritious or less-nutritious food is significantly different from zero. However, some point estimates are substantial; e.g. the effect of the relative price change for those in the tax frame to increase their purchases of nutritious food by $4.52 (relative to a mean of $36.55) and for those in the tax/subsidy frame to decrease their purchases of less nutritious food by $4.40 (relative to a mean of $49.59).

In summary, we are unable to reject the null hypothesis of no framing effect for the relative price change.

Extension: Differences by Income and Education

In our next analyses, we test whether the overall price treatment effects differed by income or education. For the sake of simplicity, we report results for expenditures (but not those for quantities) and in dollars (rather than log expenditures). Table 7 presents results of the overall price treatment effects separately for households with incomes below and above 130% of the Federal Poverty Line. For the lower-income households, the 10% relative price change was associated with households buying roughly seven dollars more each of nutritious and less-nutritious foods; however, these estimates are imprecise and not statistically significant.

Although the difference in results across income was not statistically significant, the point estimates suggest that the treatment was associated with higher-income households spending $1.27 less on nutritious food per week and $4.02 less on less-nutritious food per week, whereas lower-income households spent $7.03 more per week on nutritious food and $7.11 more per week on less-nutritious food.

Table 8 presents the results of models estimated separately by education. Again, we find no statistically significant difference between the effect of the relative price change for the two socioeconomic groups. Moreover, the difference in point estimates is considerably smaller across education groups than across income groups.

We next test whether framing effects differed by income or education. Table 9 reports results for the model that estimates treatment effects by frame, with the model estimated separately by income category. There are large and statistically significant differences in the effects of the frame by income. Specifically, low-income households that were given the subsidy frame (i.e. told that the 10% relative price change represented a subsidy for nutritious food) significantly increased their purchases of less-nutritious food (by $21.23 per week). The increase in purchases of nutritious foods was $11.58, but not statistically significant. Overall, purchases of foods rated by Guiding Stars rose $32.81 per week on average for this group.

In contrast, higher-income households that were given the subsidy frame significantly decreased their weekly purchases of both nutritious food ($4.55) and less-nutritious food ($7.55). The effects of the price change on less nutritious foods and all rated foods are significantly different for the low-income and high-income group.

Those are the differences in treatment effects across income group. In addition, *within* each income group, there is a significant difference in framing effects. In the low-income group, those given the tax frame (i.e. were told that the relative price change was a tax on less-nutritious foods) decreased their purchases of less-nutritious foods. In contrast, the low-income households given the subsidy frame bought more of everything, including less-nutritious foods. Within the higher-income group, those given the tax frame bought more nutritious food, while those given the subsidy frame bought less. Although the estimated effects of the tax frame were of opposite sign for the two income groups, the difference is not statistically significant.

Table 10 presents results for models that estimate treatment effects by frame, with the models estimated separately by education category. There are no statistically significant differences in framing effects by education. Moreover, within educational group there are no statistically significant differences in framing effects; i.e. we cannot reject the null hypothesis that the effect was the same for each treatment group or frame.

In summary, we find significant differences in framing effects by income. Specifically, those given the subsidy frame buy more food, including more of what the relative price change was seeking to discourage: less-nutritious food.

Extension: Share of Purchases that was Nutritious

As another extension, we examine the share of expenditures on rated foods that was on nutritious food. Table 11 presents results for the basic difference-in-differences model in which the dependent variable is the percent of expenditures that was on nutritious foods. The effect of the relative price change was to increase the share of expenditures devoted to nutritious food by 1.08 percentage points, relative to a mean of 42.5%. However, this increase was not statistically significant. Subsequent columns in the table list the effects for high and low income, and the high and low education groups. In no case does the relative price change result in a statistically significant change in the share of expenditures devoted to nutritious food. Yet in each case, the point estimate is small, below one percentage point.

Extension: Purchases of Unrated Foods

As described in the Data section, the Guiding Stars system rates virtually all foods in the supermarket. Those that are not rated include items that are new and have simply not yet been rated, or seasonal and therefore not consistently available. However, foods that have no calorie content are also not rated. This includes some items that are relatively uninteresting from a health perspective (e.g. dried spices) but it also includes bottled water, alcoholic beverages, and dried tea and coffee. These are of interest because after the relative price change consumers may shift away from sugar-sweetened beverages to these other drink options. In order to test for any such effects, we estimate difference-in-differences models of expenditures and quantities purchased in that category. The results appear as additional columns in each of the earlier tables. We also include a column for All Items, which includes not just rated foods but also unrated foods.

Table 5 shows that the main effect of the treatment is a very small change in weekly expenditures on unrated items ($0.81), which is not statistically significant. However, the treatment results in an increase in the quantity of unrated foods purchased per week of 0.66 units, which is statistically significant. Table 6 provides information on the effect of the framing of the relative price change. In five out of six cases, the effect of the treatment on purchases of unrated food items is not statistically significant; the exception is that those given the subsidy frame purchased 0.92 more units of unrated food per week. This effect is concentrated among the lower-income households in the subsidy frame, who increased their purchases of unrated food items by $5.78 per week (see Table 9).

Extension: Subjects' Interpretations of Relative Price Change

In order to better understand why there might be framing effects, we examine the results of a survey we administered to study participants after the treatment period ended. Participants were asked how they interpreted the treatment. Specifically, they were presented with seven statements describing the treatment, and were asked to rate their agreement with each of them on a Likert scale that ranged from 1 (strongly disagree) to 9 (strongly agree). Table 12 presents the unconditional mean responses for the entire sample as well as the control group, the entire treatment group, and each treatment group separately.

One important result that stands out is that participants, *no matter what their frame*, tended to interpret the relative price change as a subsidy for nutritious food rather than a tax on less-nutritious food. For example, for the sample as a whole, the mean agreement that the debit card payments were a "reward for eating healthy food" averaged 6.2 on the 9-point scale, whereas "penalty for eating unhealthy food" averaged 2.9. In addition, for the sample as a whole, the mean agreement that it represented a "discount for eating healthy foods" was 6.4 out of 9, whereas the agreement that it was a "tax on unhealthy foods was 3.4 out of 9.

This is not to say that the framing had no effect on subjects' perceptions. There was a statistically significant difference in the mean agreement that the treatment was a "penalty for eating unhealthy food" (3.4 in the tax frame versus 2.4 in the subsidy frame) as well as in the mean agreement that the treatment was a "tax on unhealthy foods" (3.7 in the tax frame versus 2.8 in the subsidy frame). Thus, the frame did have a detectable effect on perceptions of the treatment, but participants in all groups tended to see the treatment as more of a subsidy of nutritious food than a tax on less-nutritious food.

Extension: Subjects' Interpretations of Change in Shopping During Treatment

In the survey conducted after the treatment concluded, subjects were also asked about how being part of the study influenced their shopping. The unconditional means by group are reported in Table 12. Those in any treatment group expressed greater agreement with the statements that they bought more starred (nutritious) foods, bought healthier foods and bought a higher percentage of healthier foods, but the difference between the treatment and control groups is not statistically significant in any of those cases.

There are significant differences in the mean response to these questions by frame. Specifically, those in the tax/subsidy frame tend to express greater agreement that the study led them to buy more nutritious foods, buy healthier foods, and buy a higher percentage of healthier foods, relative to those in the subsidy frame. Notably, those differences are not reflected in the data on expenditures and quantities purchased.

Discussion

This paper contributes to the literature on the effects of food taxes and subsidies through an eight-month field experiment that created a 10% price wedge between nutritious and less-nutritious foods. We find that, on the whole, expenditures and quantities purchased did not change significantly in response to the price difference. The point estimates suggest that the treatment group bought slightly less less-nutritious food and slightly more nutritious food, but these changes were not statistically significant. Some of the point estimates are substantial in magnitude, and their lack of statistical significance is due in part to imprecision of the estimates and to limited statistical power from 208 households.

Although we hypothesized that the framing of the relative price change as either a subsidy for nutritious food or a tax on less-nutritious food could alter the treatment effect, we find no significant differences in effects by frame. We do, however, find effects of framing by income. Specifically, lower income households to whom the relative price change was framed as a subsidy bought significantly more less-nutritious food (and more of all food) than low-income households to whom it was framed as a tax. In contrast, higher-income households to whom the treatment was framed as a subsidy bought less less-nutritious food. Interestingly, survey results indicate that households in each frame tended to perceive the treatment as a discount or reward for eating healthy food.

The lower-income households may have bought more of all food, including the relatively more expensive less-nutritious food, because lower-income households may experience a large income effect of a price decrease. Previous research documented that food purchases drop significantly in the course of the benefit month for low-income households (e.g. Hastings and Washington, 2010, Shaprio, 2005) and that income increases obesity for low-income, but not other, households (see the review in Cawley, 2015). Another possibility is that poverty consumes mental bandwidth for low-income individuals (Mullainathan and Shafir, 2013) or causes distractions sufficient to cause cognitive deficits (Mani et al., 2013), such that households may have misunderstood the subsidy for nutritious food as a general "food subsidy."

Although we hypothesized that the better educated might respond differently to the treatment, we find no evidence of differences in the treatment effect or in the framing effects by education.

Taxes on energy-dense foods are arguably the most commonly-advocated anti-obesity policy. The results of this paper have several implications for such policies to promote more nutritious diets. First, taxes may need to be large to change behavior. In the U.S., taxes on soda pop and snacks average one to four percent (Chriqui et al., 2014), but we find no significant impact on expenditures or purchases from a relative price change that is 2.5 times as large as the highest existing such tax. Second, price changes may have different impacts by income; we find that subsidies for nutritious may lead low-income households to buy more of all food, including more of the less-nutritious food that the policy is attempting to discourage.

It should be noted that even if taxes do not change behavior, they can still internalize external costs, thereby addressing a market failure. Moreover, if consumers do not significantly alter their purchases, it implies that the tax results in relatively little deadweight loss and thus is a relatively efficient way for the government to collect revenue.

Strengths of this study include a randomized controlled field experiment, with actual consumers making real purchases of actual products in their usual retail environment. Such controlled field experiments represent a strong design for estimating casual effects (List, 2009). Previous research that estimated the effects of food taxes using naturally-occurring variation in prices or existing variation in taxes across states, may suffer omitted variable bias due to variation in unobserved demand or policy endogeneity (such as states with a greater preference for healthy diet being more likely to enact taxes on soft drinks or snacks). For this reason, field experiments are an advantageous method of measuring policy impacts (Harrison and List, 2004; Roe and Just, 2009; List, 2009, 2011). The present study is a relatively long experiment of this type, with a four-week baseline and 25-week treatment period.

The greatest limitation of the study is the limited statistical power associated with observing 208 households for 33 weeks; this is particularly acute when studying subsamples and testing for differences between income or education groups. In some cases, we estimate substantial point estimates but because of their imprecision they are not statistically significant. Given our limits with statistical power, we cannot rule out price elasticities common in the literature.

Readers should exercise caution when generalizing from the results associated with this relatively white, well-educated and high-income sample from upstate New York. Although we observe detailed information on food purchases, we do not observe food consumption, which would be informative about the health consequences of taxes on energy-dense foods.

Furthermore, the effects estimated in this paper may be influenced by the design of the experiment. Consumer responsiveness may have been attenuated by the fact that the price changes were less salient than usual. Our relative price changes were not reflected on supermarket shelves; consumers had to note the number of Guiding Stars for the item and take into account the subsidy or tax they received. This may have led to less responsiveness because of the mental cost of calculating the relative price change, or consumers may have overlooked the price change at times because it was less salient (Finkelstein, 2009).

In addition, participation and subsidies, minus taxes, were paid weekly, and this departure from immediacy may have also muted consumer responsiveness. Given that participants knew they were participating in a study, they may have perceived the price changes as temporary and not bothered changing their usual food habits. In addition, any announced subsidy for nutritious food or tax on less-nutritious food may have an additional effect from providing consumers with nutrition information.

In this study consumers were directed to the Guiding Stars nutrition guidance system to determine the amount of the tax or subsidy (if any). Thus, there was not only a price effect but also potentially an effect from nutrition information. This would also be true of any salient tax placed on energy-dense foods, such as a "fat tax" or tax on sugar-sweetened beverages. In contrast, it implies that the consumer responses we estimate may be greater than those that would be observed from a tax on certain foods that was implemented simply for revenue reasons and was not directly linked to the nutrition of the items.

Important directions for future research include estimating the impacts of greater price changes, testing for changes in treatment effects over time (they may increase due to habit formation or decrease due to diminishing salience or novelty), and continuing to refine how to frame price changes to maximize their intended impact.

Works Cited

Akee, Randall, Emilia Simeonova, William Copeland, Adrian Angold, and E. Jane Costello. 2013. "Young Adult Obesity and Household Income: Effects of Unconditional Cash Transfers." *American Economic Journal: Applied Economics*, 5(2): 1-28.

Brownell, K.D., Frieden, T.R., 2009. Ounces of prevention — the public policy case for taxes on sugared beverages. *New England Journal of Medicine* 360, 18.

British Medical Association. 2015. *Food for Thought: Promoting Halthy Diets among Children and Young People*.

Bureau of Labor Statistics. 2015. "Consumer Expenditures in 2013." Accessed June 17, 2015. http://www.bls.gov/cex/csxann13.pdf

Cawley, John. Forthcoming. "An Economy of Scales: A Selective Review of Obesity's Economic Causes, Consequences, and Solutions." *Journal of Health Economics*.

Centers for Disease Control and Prevention. 2015a. "Heart Disease." http://www.cdc.gov/heartdisease/faqs.htm

Centers for Disease Control and Prevention. 2015b. "Preventing Chronic Diseases: Investing Wisely in Health Preventing Diabetes and Its Complications." http://www.cdc.gov/nccdphp/publications/factsheets/Prevention/pdf/diabetes.pdf

Centers for Disease Control and Prevention. 2015c. "Addressing The Cancer Burden: At A Glance." http://www.cdc.gov/chronicdisease/resources/publications/aag/dcpc.htm

Chaloupka, Frank J., Lisa M. Powell, and Jamie F. Chriqui. 2011. "Sugar-Sweetened Beverages and Obesity: The Potential Impact of Public Policies." Journal of Policy Analysis and Management, 30(3): 645-655.

Chriqui JF, Eidson SS, Chaloupka FJ. 2014. *State Sales Taxes on Regular Soda (as of January 1, 2014) - BTG Fact Sheet*. Chicago, IL: Bridging the Gap Program, Health Policy Center, Institute for Health Research and Policy, University of Illinois at Chicago.

Elbel, Brian, Glen B. Taksler, Tod Mijanovich, Courtney B. Abrams, and L.B. Dixon. 2013. "Promotion of Healthy Eating Through Public Policy: A Controlled Experiment." Am J Prev Med 2013;45(1):49–55.

Epstein, Leonard H, Noelle Jankowiak, Chantal Nederkoorn, Hollie A Raynor, Simone A French, and Eric Finkelstein. 2012. "Experimental research on the relation between food price changes and food-purchasing patterns: a targeted review." American Journal of Clinical Nutrition, 95:789–809.

Fischer LM, Sutherland LA, Kaley LA et al. 2011. "Development and implementation of the guiding stars nutrition guidance program." *American Journal of Health Promotion* 26: e55–e63.

Fletcher, Jason M., David Frisvold, and Nathan Tefft. "Can soft drink taxes reduce population weight?." *Contemporary Economic Policy* 28, no. 1 (2010): 23-35.

Fletcher, Jason M., David E. Frisvold, and Nathan Tefft. 2011. "Are Soft Drink Taxes an Effective Mechanism for Reducing Obesity?" *Journal of Policy Analysis and Management*, 30(3): 655-662.

French SA, Story M, Jeffery RW, Snyder P, Eisenberg M, Sidebottom A, Murray D. 1997. "Pricing strategy to promote fruit and vegetable purchase in high school cafeterias." *J Am Diet Assoc*, 97:1008–10.

Gächter, Simon & Orzen, Henrik & Renner, Elke & Starmer, Chris, 2009. "Are experimental economists prone to framing effects? A natural field experiment," Journal of Economic Behavior & Organization, 70(3): 443-446.
Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." Journal of Political Economy, 80(2): 223-249.
Gruber, Jonathan H. and Sendhil Mullainathan. 2005. "Do Cigarette Taxes Make Smokers Happier," Advances in Economic Analysis and Policy, 2005, v5(1), Article 4.
Harrison, Glenn W., and John A. List. (2004). Field Experiments. Journal of Economics Literature, 42(4): 1009-1055.
Hastings, J. and Washington, E. 2010. "The first of the month effect: Consumer behavior and store responses." American Economic Journal: Economic Policy, 2(2): 142-162.
Homonoff, Tatiana A. 2015. "Can Small Incentives Have Large Effects? The Impact of Taxes versus Bonuses on Disposable Bag Use." Working paper, Cornell University.
Institute of Medicine. 2009. "Local Government Actions to Prevent Childhood Obesity." Institute of Medicine Report Brief. September 1 2009. url: http://www.iom.edu/Reports/2009/Local-Government-Actions-to-Prevent-Childhood-Obesity.aspx. Accessed Jan. 24, 2012.
Jacobson MF, Brownell KD. 2000. Small taxes on soft drinks and snack foods to promote health. American Journal of Public Health, 90(6):854–7.
Jeffery RW, French SA, Raether C, Baxter JE. 1994. "An environmental intervention to increase fruit and salad purchases in a cafeteria." Prev Med, 23:788–92.
Kahneman, Daniel, and Amos Tversky. "Prospect theory: An analysis of decision under risk." Econometrica: Journal of the Econometric Society (1979): 263-291.
Laibson, D. 1997. "Golden eggs and hyperbolic discounting." Quarterly Journal of Economics, 112(5), 443-477.
List, John A. (2009). An Introduction to Field Experiments in Economics. Journal of Economics Behavior and Organization, 70: 439-442.
List, John A. 2011. "Why Economists Should Conduct Field Experiments and 14 Tips for Pulling One Off." Journal of Economic Perspectives 25(3): 3-16.
Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. 2013. "Poverty impedes cognitive function." Science 341(6149): 976-980.
Mozaffarian, D, Jacobson MF, Greenstein JS. 2010. "Food Reformulations to Reduce Trans Fatty Acids." *New England Journal of Medicine*, 362(21): 2037-2039.
Mullainathan, Sendhil, and Eldar Shafir. Scarcity: Why having too little means so much. Macmillan, 2013.
Ogden CL, Carroll MD, Kit BK, Flegal KM. 2014. "Prevalence of Childhood and Adult Obesity in the United States, 2011-2012." JAMA, 311(8): 806-814.
Roe, Brian E., and David R. Just. 2009. "Internal and external validity in economics research: Tradeoffs between experiments, field experiments, natural experiments, and field data." American Journal of Agricultural Economics 91(5): 1266-1271.
Sassi, F., A. Belloni and C. Capobianco. 2013. "The Role of Fiscal Policies in Health Promotion", OECD Health Working Papers, No. 66, OECD Publishing.
Schmeiser, M. D. 2009. "Expanding wallets and waistlines: The impact of family income on the BMI of women and men eligible for the earned income tax credit." Health Economics, 18: 1277-1294.

Shapiro, J. M. 2005. “Is there a daily discount rate? Evidence from the food stamp nutrition cycle.” Journal of Public Economics, 89(2): 303-325.

Thow, A. M., Quested, C., Juventin, L., Kun, R., Khan, a. N., & Swinburn, B. 2011. “Taxing soft drinks in the Pacific: Implementation lessons for improving health.” *Health Promotion International*, 26(1): 55–64.

U.S. Department of Agriculture and U.S. Department of Health and Human Services. *Dietary Guidelines for Americans, 2010*. 7th Edition, Washington, DC: U.S. Government Printing Office, December 2010.

U.S. Department of Agriculture. 2013. Healthy Incentives Pilot (HIP) Interim Report, by Susan Bartlett et al. Project Officer: Danielle Berman, Alexandria, VA: July 2013.

Vesper HW, Kuiper HC, Mirel LB, Johnson CL, Pirkle JL. 2012. “Levels of Plasma trans-Fatty Acids in Non-Hispanic White Adults in the United States in 2000 and 2009.” *JAMA*, 307(6):562-563.

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

World Health Organization. 2015. *Using Price Policies to Promote Healthier Diets*. (Geneva, Switzerland: WHO).

Figure 1: Study Timeline

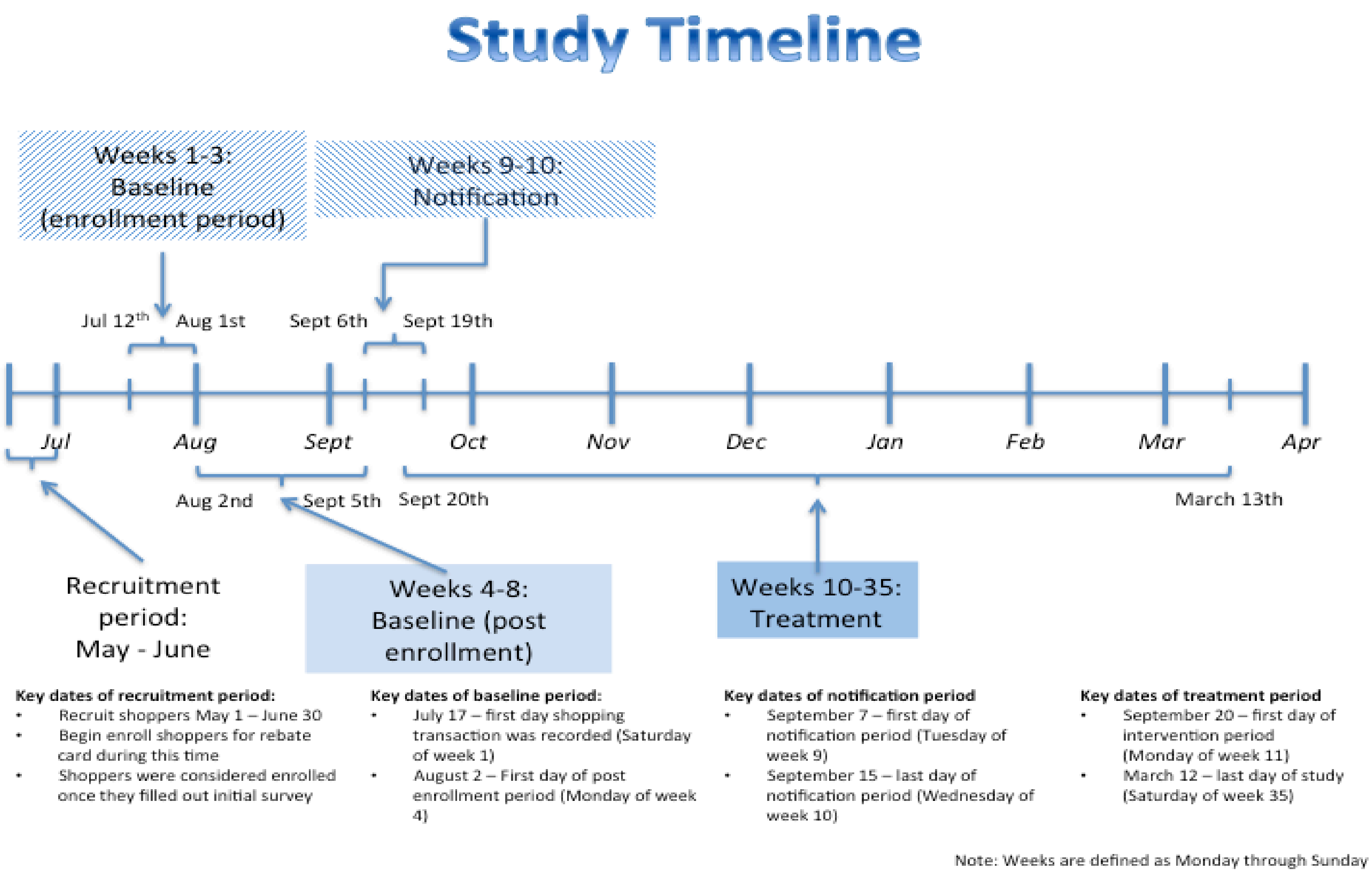

Table 1: Comparison of Treatment and Control Groups

	Control Group	Treatment Group 1: Subsidy	Treatment Group 2: Tax	Treatment Group 3: Subsidy and Tax
Discount on all Food Items as a Reward for Participation	10%	5%	15%	10%
Subsidy on Nutritious Foods	--	10%	--	5%
Tax on Less-Nutritious Foods	--	--	10%	5%
Reduction in the Relative Price of Nutritious vs Less-Nutritious Foods	None	10%	10%	10%

Table 2: Descriptive Measures of Household Demographic Variables Used in Regression (standard deviations in parentheses)

	Whole Sample	Control	All Treatment Groups	Subsidy	Tax	Tax/Subsidy
More than high school education	91.00%	92.00%	90.70%	90.60%	91.80%	89.60%
St. dev.	(0.287)	(0.274)	(0.292)	(0.295)	(0.277)	(0.309)
N (> HS ed)	182	46	136	48	45	43
N (≤ HS ed)	18	4	14	5	4	5
PIR > 1.3	81.20%	75.00%	83.20%	82.40%	82.60%	84.80%
St. dev.	(0.392)	(0.438)	(0.375)	(0.385)	(0.383)	(0.363)
N (PIR > 1.3)	155	36	119	42	48	39
N (PIR <= 1.3)	36	12	24	9	8	7
Income > $80,000	31.41%	27.08%	32.87%	25.49%	34.78%	39.13%
St. dev.	(0.465)	(0.449)	(0.471)	(0.440)	(0.482)	(0.493)
N (Inc > $80K)	60	13	47	13	16	18
N (Inc <= $80K)	131	35	96	38	30	28
More than one child under 18	58.70%	59.60%	58.40%	54.70%	56.90%	64.00%
St. dev.	(0.494)	(0.495)	(0.494)	(0.503)	(0.500)	(0.485)
N (> 1 child)	121	31	90	29	29	32
N (= 1 child)	85	21	64	24	22	18

* p<0.1. ** p<0.05. *** p<0.01. Note that the asterisks represent differences of the annotated value from the corresponding value of the control group at the respective level of significance.

Table 3: Additional Household Demographic Measures (standard errors in parentheses)

a. Food assistance, household size, and income

	Whole Sample	Control Group	All Treatment Groups	Subsidy	Tax	Tax/ Subsidy
% Households Enrolled in WIC	4.8%	5.8%	4.5%	1.8%	2.0%	10.2%
	(0.215)	(0.235)	(0.208)	(0.135)	(0.140)	(0.306)
% Households Enrolled in SNAP	4.3%	5.8%	3.9%	3.6%	3.9%	4.1%
	(0.204)	(0.235)	(0.194)	(0.189)	(0.196)	(0.200)
% Households Not Receiving Food Assistance	89.9%	87.7%	90.7%	94.4%	87.3%	89.8%
	(0.282)	(0.318)	(0.270)	(0.205)	(0.297)	(0.306)
Average Household Size	3.93	3.92	3.93	3.76	4.04	4.02
	(1.076)	(1.064)	(1.084)	(1.027)	(1.190)	(1.031)
Average Number of Children Under 18	2.2	1.8	2.3	3.0	1.9	1.8
	(3.852)	(0.936)	(4.412)	(7.295)	(1.051)	(0.889)
% Household Shopping at Hannaford	83.58	82.09	84.07	83.15	82.24	87.02
	(13.894)	(15.754)	(13.230)	(13.687)	(14.960)	(10.211)
\$10K-\$20K	9.4%	10.4%	9.0%	11.8%	4.1%	10.9%
	(0.291)	(0.309)	(0.286)	(0.325)	(0.196)	(0.315)
\$20K-\$30K	19.0%	19.5%	18.9%	19.6%	15.2%	21.7%
	(0.392)	(0.393)	(0.393)	(0.401)	(0.363)	(0.417)
\$30K-\$40K	9.7%	10.4%	9.4%	7.8%	13.0%	7.6%
	(0.294)	(0.309)	(0.290)	(0.272)	(0.341)	(0.257)
\$40K-\$50K	9.5%	12.5%	8.4%	3.9%	14.3%	7.6%
	(0.288)	(0.334)	(0.271)	(0.196)	(0.341)	(0.257)
\$50K-\$60K	12.2%	11.5%	12.4%	10.9%	13.5%	13.0%
	(0.322)	(0.314)	(0.325)	(0.303)	(0.340)	(0.341)
\$60K-\$70K	10.2%	8.3%	10.8%	12.7%	8.7%	10.9%
	(0.301)	(0.279)	(0.309)	(0.329)	(0.285)	(0.315)
\$70K-\$80K	4.9%	8.3%	3.7%	3.9%	2.8%	4.3%
	(0.213)	(0.279)	(0.186)	(0.196)	(0.153)	(0.206)
\$80K-\$90K	11.5%	10.2%	11.9%	21.6%	6.5%	6.5%
	(0.315)	(0.288)	(0.325)	(0.415)	(0.250)	(0.250)
\$90K-\$100K	4.7%	2.1%	5.5%	0.0%	8.5%	8.7%
	(0.204)	(0.144)	(0.220)	(0.000)	(0.257)	(0.285)
>\$100K	6.4%	2.6%	7.7%	5.9%	8.7%	8.7%
	(0.244)	(0.148)	(0.267)	(0.238)	(0.285)	(0.285)

* p<0.1. ** p<0.05. *** p<0.01. Note that the asterisks represent differences of the annotated value from the corresponding value of the control group at the respective level of significance.

b. Marital status and race

	Whole Sample	Control Group	All Treatment Groups	Subsidy	Tax	Tax/ Subsidy
Divorced	5.1%	8.0%	4.1%	5.7%	2.1%	4.3%
	(0.220)	(0.274)	(0.198)	(0.233)	(0.144)	(0.204)
Married	80.2%	74.0%	82.3%	77.2%	87.3%*	83.0%
	(0.381)	(0.419)	(0.366)	(0.409)	(0.297)	(0.380)
Separated	1.5%	2.0%	1.4%	1.9%	2.1%	0.0%
	(0.122)	(0.141)	(0.116)	(0.137)	(0.144)	(0.000)
Widowed	9.6%	12.0%	8.8%	9.4%	4.2%	12.8%
	(0.295)	(0.328)	(0.284)	(0.295)	(0.202)	(0.337)
Single	1.0%	0.0%	1.4%	3.8%	0.0%	0.0%
	(0.100)	(0.000)	(0.116)	(0.192)	(0.000)	(0.000)
African American	1.7%	2.0%	1.6%	1.9%	0.7%	2.1%
	(0.125)	(0.143)	(0.119)	(0.137)	(0.047)	(0.146)
American Indian or Alaska Native	0.5%	0.0%	0.7%	1.9%	0.0%	0.0%
	(0.071)	(0.000)	(0.082)	(0.137)	(0.000)	(0.000)
Asian	1.5%	2.0%	1.4%	0.0%	0.0%	4.3%
	(0.123)	(0.143)	(0.116)	(0.000)	(0.000)	(0.204)
White	93.7%	91.8%	94.3%	94.2%	94.9%	93.6%
	(0.214)	(0.236)	(0.207)	(0.208)	(0.162)	(0.247)
Hispanic or Latino	0.5%	2.0%	0.0%*	0.0%	0.0%	0.0%
	(0.071)	(0.141)	(0.000)	(0.000)	(0.000)	(0.000)
Not Hispanic or Latino	96.9%	94.0%	97.9%*	98.0%	95.6%	100.0%**
	(0.127)	(0.193)	(0.094)	(0.089)	(0.134)	(0.000)

* p<0.1. ** p<0.05. *** p<0.01. Note that the asterisks represent differences of the annotated value from the corresponding value of the control group at the respective level of significance.

Table 4: Weekly Expenditures: Unconditional Means by Treatment Group (standard deviations in parentheses)

	Whole Sample	Control Group	All Treatment Groups	Subsidy	Tax	Tax/ Subsidy
			Baseline Period			
All Foods	\$89.83	\$89.90	\$89.81	\$99.99	\$81.82	\$86.76
	(116.035)	(95.315)	(122.488)	(119.643)	(81.283)	(157.529)
All Rated Foods	\$78.80	\$78.25	\$79.00	\$88.59	\$70.25	\$77.43
	(105.460)	(83.229)	(112.223)	(113.315)	(69.960)	(143.396)
Foods Rated Less Nutritious	\$45.65	\$44.72	\$45.98	\$50.73	\$41.51	\$45.35
	(62.311)	(48.867)	(66.384)	(65.884)	(43.122)	(85.031)
Foods Rated Nutritious	\$33.15	\$33.52	\$33.02	\$37.86	\$28.74*	\$32.08
	(47.030)	(40.335)	(49.170)	(51.713)	(31.500)	(60.313)
			Treatment Period			
All Foods	\$100.88	\$101.22	\$100.76	\$109.56**	\$98.97	\$92.91**
	(102.566)	(108.558)	(100.503)	(102.659)	(97.627)	(100.332)
All Rated Foods	\$88.13	\$88.31	\$88.08	\$95.53**	\$86.33	\$81.66*
	(89.686)	(94.830)	(87.917)	(89.599)	(85.050)	(88.394)
Foods Rated Less Nutritious	\$50.65	\$51.49	\$50.37	\$54.65	\$49.37	\$46.68**
	(54.582)	(57.214)	(53.681)	(53.898)	(53.374)	(53.471)
Foods Rated Nutritious	\$37.48	\$36.82	\$37.71	\$40.88**	\$36.95	\$34.98
	(40.427)	(42.804)	(39.606)	(41.832)	(37.198)	(39.259)

Because weeks were classified as Monday through Sunday, the baseline period ended with week 8, which is the full week prior to households receiving notice of their treatment group. In the baseline period, values are set to missing prior to the first shopping trip in the first three weeks. Once all households were enrolled in the study (by week four), any missing value was set to zero. Since households received their notices between September 7-15, weeks including these dates were omitted from the analysis. As a result, the treatment period begins with week 11, which is after all households received notice of their treatment.

* $p<0.1$. ** $p<0.05$. *** $p<0.01$. Note that the asterisks represent differences of the annotated value from the corresponding value of the control group at the respective level of significance.

Table 5: Overall Price Effect on Weekly Household Expenditures and Quantities Purchased (standard errors in parentheses)

	Expenditures					Quantities				
	Nutritious	Less Nutritious	All Rated Items	Unrated	All Items	Nutritious	Less Nutritious	All Rated Items	Unrated	All Items
All Treatment Groups	$1.11	-$1.55	-$0.44	$0.81	$0.37	0.951	-0.873	0.078	0.661*	0.739
	(3.010)	(4.042)	(6.780)	(1.138)	(7.606)	(1.347)	(1.607)	(2.822)	(0.387)	(3.091)
Weekly Dummy Variables	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
N	6572	6572	6572	6572	6572	6572	6572	6572	6572	6572
Unconditional mean of dependent variable	$36.55	$49.59	$86.14	$11.86	$98.50	16.132	18.853	34.985	3.609	38.744

Participants in the intervention conditions were all combined. Regression coefficients were estimated using a fixed effects regression with weekly dummy variables. For the sake of space, coefficients from the weekly dummy variables were not included in the table. Because weeks were classified as Monday through Sunday, the baseline period ended with week 8, which is the full week prior to households receiving notice of their treatment group. In the baseline period, values are set to missing prior to the first shopping trip in the first three weeks. Once all households were enrolled in the study (by week four), any missing value was set to zero. Since households received their notices between September 7-15, weeks including these dates were omitted from the analysis. As a result, the treatment period begins with week 11, which is after all households received notice of their treatment.
* p<0.1. ** p<0.05. *** p<0.01

Table 6: Impact of Price Frame on Expenditures and Quantities Purchased (standard errors in parentheses)

	Expenditures					Quantities				
	Nutritious	Less Nutritious	All Rated Items	Unrated	All Items	Nutriti ous	Less Nutritious	All Rated Items	Unrated	All Items
Subsidy	-$0.78	-$2.29	-$3.07	$1.60	-$1.47	0.523	-1.220	-0.698	0.917**	0.220
	(3.655)	(4.914)	(8.225)	(1.376)	(9.041)	(1.600)	(1.884)	(3.327)	(0.450)	(3.627)
Tax	$4.52	$1.89	$6.41	-$0.07	$6.34	2.287	0.896	3.182	0.306	3.489
	(3.489)	(4.784)	(7.908)	(1.460)	(9.015)	(1.564)	(1.925)	(3.325)	(0.461)	(3.654)
Tax/Subsidy	-$0.42	-$4.40	-$4.82	$0.84	-$3.98	-0.002	-2.384	-2.386	0.752	-1.634
	(4.371)	(5.831)	(9.942)	(1.466)	(11.010)	(1.876)	(2.293)	(4.044)	(0.527)	(4.399)
Weekly Dummy Variables	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
N	6572	6572	6572	6572	6572	6572	6572	6572	6572	6572
Unconditional Mean of Dependent Variable	$36.55	$49.59	$86.14	$11.86	$98.50	16.132	18.853	34.985	3.609	38.744

Participants in the intervention conditions were all combined. Regression coefficients were estimated using a fixed effects regression with weekly dummy variables. For the sake of space, coefficients from the weekly dummy variables were not included in the table. Because weeks were classified as Monday through Sunday, the baseline period ended with week 8, which is the full week prior to households receiving notice of their treatment group. In the baseline period, values are set to missing prior to the first shopping trip in the first three weeks. Once all households were enrolled in the study (by week four), any missing value was set to zero. Since households received their notices between September 7-15, weeks including these dates were omitted from the analysis. As a result, the treatment period begins with week 11, which is after all households received notice of their treatment.

* p<0.1. ** p<0.05. *** p<0.01

a. p<0.05 for difference between Subsidy and Tax

b. p<0.05 for difference between Subsidy and Tax/Subsidy

c. p<0.05 for difference between Tax and Tax/Subsidy

Table 7: Overall Price Effect on Weekly Household Expenditures when Separated by PIR = 1.3 (standard errors in parentheses)

	Poverty Income Ratio <= 1.3					Poverty Income Ratio > 1.3				
	Nutritious	Less Nutritious	All Rated Items	Unrated	All Items	Nutritious	Less Nutritious	All Rated Items	Unrated	All Items
All Treatment Groups	$7.03	$7.11	$14.14	$2.47	$16.61	-$1.27	-$4.02	-$5.29	$0.24	-$5.05
	(6.010)	(9.793)	(15.460)	(2.597)	(17.420)	(3.707)	(4.543)	(7.898)	(1.313)	(8.893)
Weekly Dummy Variables	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
N	1141	1141	1141	1141	1141	4904	4904	4904	4904	4904
Unconditional Mean of Dependent Variable	$28.28	$41.04	$69.32	$9.17	$78.85	$38.36	$50.70	$89.06	$12.25	$101.81

Participants in the intervention conditions were all combined. Regression coefficients were estimated using a fixed effects regression with weekly dummy variables. For the sake of space, coefficients from the weekly dummy variables were not included in the table. Because weeks were classified as Monday through Sunday, the baseline period ended with week 8, which is the full week prior to households receiving notice of their treatment group. In the baseline period, values are set to missing prior to the first shopping trip in the first three weeks. Once all households were enrolled in the study (by week four), any missing value was set to zero. Since households received their notices between September 7-15, weeks including these dates were omitted from the analysis. As a result, the treatment period begins with week 11, which is after all households received notice of their treatment.

* p<0.1. ** p<0.05. *** p<0.01.

d. p<0.05 difference of estimates for the same type of food (all items, all rated items, etc.) but across demographic comparisons.

Table 8: Overall Price Effect on Weekly Household Expenditures when Separated by Education Level (standard errors in parentheses)

	High School Education or Less					More than High School Education				
	Nutritious	Less Nutritious	All Rated Items	Unrated	All Items	Nutritious	Less Nutritious	All Rated Items	Unrated	All Items
All Treatment Groups	$2.36	-$4.02	-$1.65	$6.18	$4.52	$0.52	-$2.17	-$1.65	$0.46	-$1.19
	(11.190)	(20.950)	(31.600)	(4.130)	(34.200)	(3.091)	(3.925)	(6.714)	(1.139)	(7.528)
Weekly Dummy Variables	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
N	567	567	567	567	567	5759	5759	5759	5759	5759
Unconditional Mean of Dependent Variable	$25.16	$39.92	$65.08	$8.76	$74.23	$37.73	$50.41	$88.14	$12.05	$100.67

Participants in the intervention conditions were all combined. Regression coefficients were estimated using a fixed effects regression with weekly dummy variables. For the sake of space, coefficients from the weekly dummy variables were not included in the table. Because weeks were classified as Monday through Sunday, the baseline period ended with week 8, which is the full week prior to households receiving notice of their treatment group. In the baseline period, values are set to missing prior to the first shopping trip in the first three weeks. Once all households were enrolled in the study (by week four), any missing value was set to zero. Since households received their notices between September 7-15, weeks including these dates were omitted from the analysis. As a result, the treatment period begins with week 11, which is after all households received notice of their treatment.

* $p<0.1$. ** $p<0.05$. *** $p<0.01$.

d. $p<0.05$ difference of estimates for the same type of food (all items, all rated items, etc.) but across demographic comparisons.

Table 9: Impact of Price Frames on Weekly Expenditures When Separated by PIR = 1.3 (standard errors in parentheses)

	Poverty Income Ratio <= 1.3					Poverty Income Ratio > 1.3				
	Nutritious	Less Nutritious	All Rated Items	Unrated	All Items	Nutritious	Less Nutritious	All Rated Items	Unrated	All Items
Subsidy	11.58	$21.23* [a]	$32.81* [ad]	$5.78** [a]	$38.59** [ad]	-4.548 [a]	-7.546 [d]	-12.09 [ad]	0.414	-11.68 [d]
	(6.914)	(10.780)	(16.990)	(2.802)	(18.990)	(4.434)	(5.521)	(9.534)	(1.608)	(10.490)
Tax	$0.30	-$9.037 [a]	-$8.735 [a]	-$3.38 [a]	-$12.11 [a]	3.832 [a]	3.62	7.451 [a]	0.588	8.039
	(8.190)	(12.470)	(20.380)	(4.138)	(23.370)	(4.180)	(5.334)	(9.015)	(1.540)	(10.230)
Tax/Subsidy	$9.14	$8.14	$17.28	$5.13**	$22.40	-2.831	-7.931	-10.76	-0.327	-11.09
	(6.874)	(9.965)	(16.310)	(2.039)	(17.710)	(5.338)	(6.790)	(11.800)	(1.750)	(13.080)
Weekly Dummy Variables	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
N	1141	1141	1141	1141	1141	4904	4904	4904	4904	4904
Unconditional means of dependent variables										
Mean	$28.28	$41.04	$69.32	$9.17	$78.85	$38.36	$50.70	$89.06	$12.25	$101.81

Participants in the intervention conditions were all combined. Regression coefficients were estimated using a fixed effects regression with weekly dummy variables. For the sake of space, coefficients from the weekly dummy variables were not included in the table. Because weeks were classified as Monday through Sunday, the baseline period ended with week 8, which is the full week prior to households receiving notice of their treatment group. In the baseline period, values are set to missing prior to the first shopping trip in the first three weeks. Once all households were enrolled in the study (by week four), any missing value was set to zero. Since households received their notices between September 7-15, weeks including these dates were omitted from the analysis. As a result, the treatment period begins with week 11, which is after all households received notice of their treatment.

* p<0.1. ** p<0.05. *** p<0.01.
a. p<0.05 for difference between Subsidy and Tax
b. p<0.05 for difference between Subsidy and Tax/Subsidy
c. p<0.05 for difference between Tax and Tax/Subsidy
d. p<0.05 difference of estimates for the same type of food (all items, all rated items, etc.) but across demographic comparisons.

Table 10: Impact of Price Frame on Weekly Expenditures When Separated by Education (standard errors in parentheses)

	High School Education or Less					More than High School Education				
	Nutritious	Less Nutritious	All Rated Items	Unrated	All Items	Nutritious	Less Nutritious	All Rated Items	Unrated	All Items
Subsidy	-$0.65	-$3.86	-$4.51	$7.38	$2.87	-$0.97	-$2.71	-$3.68	$1.37	-$2.31
	(11.440)	(21.320)	(32.150)	(6.381)	(34.410)	(3.824)	(4.986)	(8.414)	(1.343)	(9.226)
Tax	$2.26	-$5.53	-$3.26	$6.79*	$3.53	$4.19	$2.34	$6.53	-$0.44	$6.09
	(12.020)	(23.800)	(34.630)	(3.621)	(36.810)	(3.536)	(4.636)	(7.781)	(1.523)	(8.963)
Tax/Subsidy	$5.64	-$2.81	$2.83	$4.35	$7.17	-$1.81	-$6.52	-$8.33	$0.39	-$7.94
	(13.210)	(24.060)	(36.560)	(4.385)	(39.780)	(4.705)	(5.961)	(10.400)	(1.533)	(11.470)
Weekly Dummy Variables	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
N	567	567	567	567	567	5759	5759	5759	5759	5759
Unconditional Mean of Dependent Variable	$25.16	$39.92	$65.08	$8.76	$74.23	$37.73	$50.41	$88.14	$12.05	$100.67

Participants in the intervention conditions were all combined. Regression coefficients were estimated using a fixed effects regression with weekly dummy variables. For the sake of space, coefficients from the weekly dummy variables were not included in the table. Because weeks were classified as Monday through Sunday, the baseline period ended with week 8, which is the full week prior to households receiving notice of their treatment group. In the baseline period, values are set to missing prior to the first shopping trip in the first three weeks. Once all households were enrolled in the study (by week four), any missing value was set to zero. Since households received their notices between September 7-15, weeks including these dates were omitted from the analysis. As a result, the treatment period begins with week 11, which is after all households received notice of their treatment.

* $p<0.1$. ** $p<0.05$. *** $p<0.01$.

a. $p<0.05$ for difference between Subsidy and Tax

b. $p<0.05$ for difference between Subsidy and Tax/Subsidy

c. $p<0.05$ for difference between Tax and Tax/Subsidy

d. $p<0.05$ difference of estimates for the same type of food (all items, all rated items, etc.) but across demographic comparisons.

Table 11: Overall Price Effect on Shares of Expenditures on Nutritious Foods, by PIR and Income Level (standard errors in parentheses)

	All	PIR <= 1.3	PIR > 1.3	HS Educ or Less	More than HS Educ
All Treatments	0.0108	0.00359	0.00834	-0.0057	0.00928
	(0.01)	(0.03)	(0.01)	(0.03)	(0.01)
Weekly Dummy Variables	✔	✔	✔	✔	✔
N	4816	769	3637	342	4266
Unconditional Mean Shares	0.425	0.406	0.433	0.369	0.431

Shares of less nutritious and nutritious foods were calculated using only rated food purchases, thus the sign of the share is opposite when comparing nutritious and less nutritious foods. Participants in the intervention conditions were all combined. Regression coefficients were estimated using a fixed effects regression with weekly dummy variables. For the sake of space, coefficients for the constants and the weekly dummy variables were not included in the table. Because weeks were classified as Monday through Sunday, the baseline period ended with week 8, which is the full week prior to households receiving notice of their treatment group. In the baseline period, values are set to missing prior to the first shopping trip in the first three weeks. Once all households were enrolled in the study (by week four), any missing value was set to zero. Since households received their notices between September 7-15, weeks including these dates were omitted from the analysis. As a result, the treatment period begins with week 11, which is after all households received notice of their treatment.

* p<0.1. ** p<0.05. *** p<0.01.

d. p<0.05 difference of estimates for the same type of food (all items, all rated items, etc.) but across demographic comparisons.

Table 12: Results of Post-Experiment Survey (on 9-point Likert Scale)

	Whole Sample	Control Group	All Treatment Groups	Subsidy	Tax	Tax/ Subsidy
Interpretation of Treatment:						
Penalty for eating unhealthy food	2.9	2.6	3.0	2.4[a]	3.4[a]	3.2
	(1.937)	(1.739)	(2.003)	(1.662)	(2.100)	(2.161)
Reward for eating healthy food	6.2	6.1	6.3	6.0	6.0	6.9
	(2.286)	(2.515)	(2.211)	(2.362)	(2.394)	(1.641)
Tax on unhealthy foods	3.4	2.8	3.6*	2.8[b]	3.7*	4.4**[b]
	(2.076)	(1.796)	(2.141)	(1.696)	(2.237)	(2.218)
Discount for eating healthy foods	6.4	5.8	6.6*	6.7	6.2	6.9*
	(2.225)	(2.543)	(2.077)	(2.157)	(2.313)	(1.595)
Effective in changing what I usually buy	4.5	4.2	4.6	4.8	4.2	5.0
	(2.419)	(2.444)	(2.413)	(2.250)	(2.452)	(2.568)
How much did being a part of the study influence your shopping?						
Buy more starred foods	5.0	4.5	5.1	4.8[b]	4.8[c]	5.9[bc]
	(2.084)	(2.152)	(2.048)	(2.009)	(2.060)	(1.950)
Buy more non-starred foods	3.1	3.2	3.1	3.0	3.2	3.0
	(1.421)	(1.567)	(1.373)	(1.650)	(1.050)	(1.401)
Buy healthier food	5.3	4.7	5.5	5.0[b]	5.3	6.2[b]
	(2.146)	(2.271)	(2.078)	(2.048)	(2.357)	(1.541)
Buy a higher percentage of healthy food	5.3	4.8	5.5	4.9[b]	5.5	6.2[b]
	(2.200)	(2.360)	(2.124)	(2.043)	(2.407)	(1.595)
In general, over the entire program:						
Shopped healthier at the beginning than at the end	3.3	3.1	3.4	3.4	3.1	3.6
	(1.725)	(1.555)	(1.784)	(1.845)	(1.465)	(2.077)

Note that the asterisks represent differences of the annotated value from the corresponding value of the control group at the respective level of significance. All responses were based on a 9 point Likert scale from Strongly Disagree (1) to Strongly Agree (9).
* $p<0.1$. ** $p<0.05$. *** $p<0.01$.
a. $p < 0.05$ for comparison between Subsidy and Tax groups.
b. $p< 0.05$ for comparison between Subsidy and Tax/Subsidy groups.
c. $p<0.05$ for comparison between Tax and Tax/Subsidy groups.