The Healthy Incentives Pilot and Fruit and Vegetable Intake: Interim Results

Jacob Alex Klerman
Susan Bartlett
Parke Wilde
Lauren Olsho


*AAEA member submission

Copyright 2013 by [authors]. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.
The Healthy Incentives Pilot and Fruit and Vegetable Intake: Interim Results

Jacob Alex Klerman
Susan Bartlett
Parke Wilde
Lauren Olsho

Corresponding author:
Jacob Alex Klerman
Abt Associates Inc.
55 Wheeler St., Cambridge, MA 02138
Phone: (617) 520-2613
Fax: (617) 386-8491
Jacob_Klerman@abtassoc.com
Acknowledgements: This paper is based on analyses performed by Abt Associates under contract to the U.S. Department of Agriculture Food and Nutrition Service as part of the official evaluation of the Health Incentives Pilot (Contract number AG-3198-D-10-0044 with the Food and Nutrition Service, United States Department of Agriculture). That evaluation would not have been possible without the cooperation of USDA FNS staff including Danielle Berman, Anita Singh, Kelly Kinnison, Usha Kalro, and Kathryn Law; and Massachusetts Department of Transitional Assistance HIP director Frank Martinez Nocito. At Abt, Marianne Beauregard, Michelle Blocklin, Patty Connor, Ayesha Enver, Chris Hamilton, Don LaLiberty, Chris Logan and Patricia McKinney contributed to the performance of the evaluation. Under the leadership of Susie McNutt and Crystal MacAllum, Westat conducted the household survey.

No HIP Evaluation funds were used in writing this paper.

This paper has benefitted from helpful comments from the participants at the Abt Journal Author Support Group (in particular Bill Rhodes) and Cindy Taylor.
THE HEALTHY INCENTIVES PILOT AND FRUIT AND VEGETABLE INTAKE: INTERIM RESULTS

Abstract

In response to low consumption of fruits and vegetables by SNAP recipients, the USDA Food and Nutrition Service created the Healthy Incentives Pilot (HIP) to test the efficacy of providing a 30% incentive for purchases of fruits and vegetables. Published elasticity estimates imply that a pure price reduction of 30% would increase fruit and vegetable consumption by about 20%, i.e., about a fifth of a cup per day. This paper considers the applicability of predictions based on a pure price reduction. It then reports interim results of a random assignment evaluation of HIP which find an increase of about a fifth of a cup per day.
American fruit and vegetable consumption is well below levels specified in federal dietary guidelines, and the discrepancy is particularly large among poorer Americans (US DHHS 2010; Guthrie et al. 2007; Cole and Fox, 2008; Briefel 2007). Following on analyses by Frazao et al. (2007), Guthrie et al. (2007), and Dong and Lin (2009), the Food, Conservation, and Energy Act of 2008 included authorized funds for pilot projects to determine if financial incentives provided to SNAP recipients at the point of sale increase the consumption of fruits, vegetables, or other healthful foods. On the basis of this legislative authority, USDA’s Food and Nutrition Service (FNS) designed the Healthy Incentives Pilot (HIP) which allowed Supplementary Nutrition Assistance Program (SNAP) households to earn a 30% incentive on targeted fruit and vegetable (TFV) purchases; i.e., a 30% decrease in the net price of TFVs that are purchased with SNAP benefits at participating retailers. (TFVs are defined below.)

The legislation authorizing HIP also required a rigorous evaluation of its impacts. The research design was based on random assignment. Data collection included three rounds of participant surveys: before implementation of the pilot (Round 1), 4-6 months after implementation (Round 2), and 9-11 months after implementation (Round 3). This paper reports interim findings from that evaluation (based on data through Round 2 only): HIP induces an increase in self-reported TFV intake of about a fifth of a cup (0.215 cups on a base of 0.878 cups; i.e., about 24.5%). However, even with HIP, fruit and vegetable intake remains well below recommended levels for adults, which range from 3.5 to 5.0 cups depending on age, sex, and level of physical activity.

The balance of this paper proceeds as follows. The first section describe HIP; the second section projects the likely impact of the price incentive based on pre-HIP
evidence. The third section describes the evaluation, including randomization and data collection. The fourth section reports the experimental estimates of impact and relates them back to the likely impact estimates estimated from the literature. The final section discusses the implications of the findings and the remaining evaluation.

**The Healthy Incentives Pilot**

Various approaches have been proposed to increase fruit and vegetable consumption among adults. Most prominent are nutrition education and financial incentives (Guthrie et al. 2007; GAO 2008). HIP implemented a financial incentive in which participating SNAP households received a 30% incentive for TFV purchases (defined below). Specifically, when a household made a TFV purchase in a participating retailer using its SNAP electronic benefit transfer (EBT) card, a 30% incentive was credited to the household’s EBT account. Incentives earned were capped at $60, but in practice that cap was reached only rarely—in well under 1% of household months.

Testing such a monetary incentive for fruit and vegetable consumption in SNAP required a local, rather than national, implementation and evaluation. Implementation of the incentive for the pilot required retailer cooperation. In a national pilot, participating SNAP households would be too rare to induce retailer participation; in a smaller geographical area the fraction of customers who would be in the pilot would be larger, making retailer participation more attractive (7,500 participating households out of more than 20 million SNAP households nationwide). Therefore, FNS issued a request for proposals, reviewed bids, and awarded the implementation contract to the Massachusetts Department of Transitional Assistance (DTA) to operate the pilot in Hampden County—
State officials then proceeded to recruit retailers to participate in the Pilot and funded the EBT provider to make appropriate changes to its software. In practice, retailer participation required either modifying existing integrated electronic cash register (IECR) systems or manually separating TFV purchases from other SNAP purchases. From a retailer perspective, the advantages of making these changes were unclear. The 7,500 pilot households included 17,000 individuals. They represent less than 4% of Hampden County’s population and the pilot was only to last 15 months. Retailers needed to make moderate investments (for the most part, paid for by the pilot) to change their IECR systems or to train staff, but only for a small population and short time period, minimizing potential direct benefits for retailers. Consistent with this observation, retailer recruitment was a challenge. Measured by SNAP benefits redeemed, about 60% of all retailers participated. In particular, one of the large retail chains in Hampden County did not participate. (Again, see Bartlett, Beauregard et al. [2013].)

HIP incentives were earned only for TFV purchases in participating retailers using SNAP EBT benefits. With appropriate modifications to point-of-sale (POS) systems, retailers with IECR systems automatically credited the incentive to EBT accounts as part of their processing. In retailers without IECR systems (i.e., most non-supermarkets/superstores), a manual process for crediting incentives was required. In particular, the HIP participant needed to request that the TFVs be rung up separately from other SNAP purchases and the retailer needed to comply.
TFVs that qualified for the incentive could be fresh, canned, frozen, or dried fruits and vegetables, but they could not include added sugars, fats, or oils, and fruits could not include added salt. Also excluded were white potatoes and 100% fruit juice. The TFVs are the same fruits and vegetables that are eligible for the WIC Fruit and Vegetable Cash Value Voucher. Finally, the SNAP benefit cannot be used to purchase food at restaurants, so the HIP incentive could not be earned for those purchases.

What Impact Would be Plausible?

This section discusses the plausible impact of HIP. The discussion begins with a standard elasticity-based discussion of HIP’s financial incentive and then discusses in turn five considerations suggesting that the simple standard elasticity-based analysis is incomplete: (i) there is ambiguity about the appropriate true elasticity; (ii) the projections assume that households will spend any incentives earned as though they were cash; (iii) incentives are not earned on all TFV purchases; (iv) households may not perfectly understand the program; and (v) HIP includes implicit marketing components.

Elasticity-Based Analysis

In pure economic terms, HIP impacts could occur through two separate mechanisms. First, TFV purchases would earn the incentive, increasing total SNAP benefits and thereby inducing an income effect. Second, the net price of incremental TFV purchases would fall by 30% (i.e., the size of the HIP incentive).

We can project the likely impact of a discrete change in the price of TFVs using the Slutsky equation:

$$\varepsilon_{TFV} = \varepsilon_{TFV}^h - s_{TFV} \eta_{TFV}$$
where $\varepsilon_{TFV}$ is the uncompensated price elasticity of TFVs (i.e., the percentage change in purchases with a 1% change in price); $\varepsilon_{TFV}^h$ is the compensated elasticity of TFVs (i.e., the percentage change in purchases with a 1% change in price, holding utility constant); $s_{TFV}$ is the share of TFVs; and $\eta_{TFV}$ is the income elasticity of TFVs (i.e., the percentage change in fruit and vegetable consumption with a percentage change in income).

We can roughly approximate each of these terms.

- **Share of TFVs**: From the published 2011 Consumer Expenditure Survey (CEX) tables (see [http://www.bls.gov/cex/csxstnd.htm](http://www.bls.gov/cex/csxstnd.htm)) for the lowest income quintile, the share of TFVs in total expenditure is roughly 1.6%.

- **Income Elasticity of TFVs**: Also from the 2011 CEX, we can roughly approximate the income elasticity of TFVs. Between the first and second quintiles, household income increases from $22,001 to $32,092; i.e., 45%. Between those same quintiles, fruit and vegetable consumption increases from $448 to $556; i.e., 24%. Thus, a rough estimate of the income elasticity can be computed as the ratio of these two figures: $0.53 = 24\%/45\%$. These estimates are roughly consistent with other studies of the income elasticity of fruits and vegetables (e.g., Lin 2005; Blisard, Steward, and Jolliffe 2004; Frazao et al. 2007).

- **Compensated Price Elasticity**. Based on the reviewed literature, a compensated elasticity of two-thirds is plausible (and computationally convenient). Specifically, Andreyeva, Long, and Brownell (2010) surveyed the literature on fruit and vegetable elasticities. They identified 20 estimates of the elasticity of fruit purchases, with a mean of 0.70, and a range of 0.16 to
3.02 across studies. They also identified 20 estimates of the elasticity of vegetable purchases, with a mean of 0.58 and a range of 0.20 to 1.11. Using the Slutsky equation, we can convert this into an uncompensated elasticity. Given the income elasticity of 0.53 and the trivial 2% income share, the compensated elasticity is about 0.01 less than the uncompensated elasticity.

A naive extrapolation to HIP might proceed as follows. Following on Andreyeva, Long, and Brownell’s work, the simple average of the mean elasticity estimates for fruits and vegetables is very close to two-thirds (i.e., 0.64). HIP provides a 30% incentive for TFV purchases, so in rounded terms we would expect HIP to increase TFV spending by roughly 20%.

The implied income effects are small. Baseline TFV intake is about one cup per adult per day, so a 20% increase in consumption is equivalent to about a fifth of a cup.

The balance of this section discusses five caveats to this simple elasticity analysis.

*Ambiguity about the True Appropriate Elasticity*

There is considerable ambiguity about the appropriate elasticity. First, each of the existing estimates has considerable sampling variability (i.e., large confidence intervals).

Second, the existing studies apply to different populations, and none of those populations are exactly comparable to TFVs purchased by SNAP participants in Hampden County in 2012. It should be noted, however, that Andreyeva, Long, and Brownell’s (2010) review does not suggest strong evidence of heterogeneity of elasticities.

Third, the sources of variation exploited in the literature vary across studies. Almost all of the variation is observational. It is not clear whether observed differences in
prices represent true variation in prices faced rather than choices of where to shop and regional differences that might be correlated with other factors (e.g., tastes for fruits and vegetables). Thus, it is not clear that such observational studies estimate the true causal effect of an exogenous change in the price of fruits and vegetables (as is required to compare to HIP).

Finally, and in part as a result of the considerations just discussed, there is considerable variation across studies in the elasticities. The computations above use a simple average—across studies and across estimates for fruits and estimates for vegetables. In conclusion and in support of these estimates, note that the elasticities used above are consistent with Dong and Lin’s (2009) literature review. They are also consistent with Dong and Lin’s own estimates from Neilsen Homescan Consumer Panel data: 0.52 for fruits and 0.69 for vegetables, with no significant difference between low income and other households.

**Will Households Spend the Incentive like Cash?**

These projections implicitly assume that households can and will spend the incentive like cash. In SNAP, the Southworth Hypothesis suggests that in-kind SNAP benefits are equivalent to cash for inframarginal households. The projections implicitly assume no “stickiness” of the incentive (i.e., that households do not feel some obligation to spend the incentive on food) and that households are not otherwise constrained (i.e., that they otherwise spend cash on food). These conditions seem plausible: (i) as already noted, the income effect is small; (ii) the SNAP benefit formula assumes that households with income will spend 30% of that income on food; and (iii) estimates of the Southworth Parameter suggest that only about 20% of incremental SNAP benefits are spent on food
(Fox, Hamilton, and Lin 2004; see Hanson 2010 for a more recent discussion consistent with this estimate; but see Klerman [2013] for a larger estimate).

**Incentives Not Earned on All TFV Purchases**

Crucially, these projections implicitly assume that the decrease in price is uniform. HIP only applies to SNAP purchases, so the net price does not decline for cash purchases. Thus, households that currently buy some TFVs with cash would need to switch TFV purchases from cash to SNAP and other SNAP-qualifying purchases from SNAP to cash. This seems plausible, if the household understands how the incentive is earned.

Furthermore, the net price only declined at participating retailers, who as noted above represent only about 60% of all retailers (by SNAP purchases). At one extreme, if the cost of changing retailers was very high (e.g., in food deserts; see Ver Ploeg et al. 2009), we might expect only households that shopped at participating retailers to change TFV purchases; therefore, the impact would be about half of that projected above. However, HIP gives households an incentive to switch to participating retailers. Thus, we would expect the impact to be greater than the no switching benchmark. For households that already shop in multiple retailers or who have a participating retailer easily accessible, we might expect switching of retailers. For those for whom the participating retailer is much less convenient than the non-participating retailer, switching seems less likely. The incentive may not have been sufficient to compensate for the out of pocket and time costs of switching (or adding) a retailer.

**Imperfect Understanding of HIP**

These elasticities are estimated for true price differences. HIP only provides a subjective net price decrease if participants understand the incentive. The cognitive task required
(e.g., computing percentage discounts) is non-trivial and evidence from other changes to
the budget set suggest that understanding is likely to be incomplete (Meyers et al. 1998).

DTA conducted intensive outreach to inform those assigned to HIP about the
program. That outreach included multiple mailings. DTA also scheduled multiple in-
person training sessions, but attendance was very low—approximately 100 people,
representing 1.3% of HIP households. Similarly, DTA set up a help line, but it only
received about 500 calls. Assuming no multiple calls by a given household, this would be
6.7% of HIP households.

Some forms of incomplete understanding would lead to lower impacts. For
example, people who are unaware of the program or totally do not understand what a
30% incentive means would probably have zero impacts. In fact, 39% of those assigned
to HIP reported not having heard of HIP at the Round 2 interview; 62% of those who had
heard of HIP reported they had seen a letter. No other source of information was
mentioned by more than 12% of those assigned to HIP and many of the responses seem
inconsistent with plausible sources of information about HIP (5% word of mouth; 2%
media).

Other forms of incomplete understanding might lead to larger impacts.
Specifically, it is possible that people believed that the program was broader than it was.
For example, below we show some evidence suggesting that some people might have
thought that fruit juices qualified for the incentive. In addition, people might have
thought that they qualified whenever they used their SNAP card. Some people might
have thought that they could earn the incentive whenever they used their SNAP card to
purchase TFVs; i.e., they may not have realized that some retailers were not participating
in HIP. Some people might have thought that they could earn the incentive with TFV purchases using cash.

Implicit Marketing

Last, and suggesting that the impacts might be larger than projected based on the elasticities, HIP was not a simple net price decrease. HIP households received a variety of HIP-related materials. Those materials began with three mailings in the three weeks before the household’s HIP start date. Each mailing included a cover letter and successively more training materials (a brochure, guidelines as to which fruits and vegetables were HIP-eligible, a list of participating retailers, answers to frequently asked questions, and a HIP-specific “sleeve” for the EBT card). Throughout the pilot, DTA sent out additional mailings, including another brochure and updated lists of participating retailers and farmers markets.

In addition to providing information on the program, these materials were intended to work as some combination of marketing of TFVs. Existing empirical evidence and the standard practice of marketers would suggest that a price cut combined with marketing might have larger impacts than the price cut alone (see Dong and Leibtag 2010).

Consistent with this marketing caveat, HIP households are significantly more likely than non-HIP households to report having heard or seen messages about fruits and vegetables (76.6% vs. 69.6%; p<0.01). However, there was no evidence of difference in attitudes with respect to fruits and vegetables (enjoy trying new fruits, enjoy trying new vegetables, encourage family and friends to eat fruits and vegetables). Finally, there was
no difference in responses to a question about not buying fruits and vegetables because they cost too much.

**Discussion of Likely Effects of HIP Incentive Based on Previous Literature**

In summary, the most plausible estimate is probably below the simple elasticity computation of a fifth of a cup. By how much is unclear. Considerations of incomplete understanding and incomplete retailer participation imply that estimates half that size (i.e., a tenth of a cup) would not have been implausible. On the other hand, considering marketing, impacts might have been larger than is implied by this augmented elasticity analysis. Finally, uncertainty about the value of the appropriate true elasticity implies that there is considerable ambiguity about the estimated impact. Thus, while we can project a likely impact of HIP based on elasticities, there is reason to doubt how appropriate those projections would be for HIP as implemented.

**The Evaluation**

HIP was offered to a randomly selected sample of 7,500 households; i.e., about 14% of the approximately 55,000 SNAP households in Hampden County. The other households received standard benefits under the regular SNAP rules (i.e., they were not eligible for incentives). This section provides further detail on four aspects of the evaluation: (i) follow-up rules; (ii) data collection; (iii) measuring TFV intake; and (iv) sample statistics.

**Follow-Up Rules**

When properly implemented, a conventional random assignment design eliminates selection bias; i.e., at randomization the HIP group and the non-HIP group do not differ systematically. Thus, HIP/non-HIP comparisons can be interpreted as the causal impact of HIP.
One of the characteristics of a conventional random assignment study is that everyone randomized is followed (i.e., follow-up data is collected and analyzed). However, the HIP evaluation did not follow everyone who was randomized. Specifically, those who left SNAP between randomization and the first follow-up survey were not interviewed. Evaluation costs are driven by the size of the survey sample. For the given evaluation budget, the decision not to follow those who leave SNAP allows a larger sample of those on HIP, maximizing statistical power. This strategy seemed attractive given that any impacts on those who leave SNAP were likely to be second-order. They have no change in net price; the only effect of being assigned to HIP would be some potential change in food patterns due to (short-term) exposure to HIP.

This follow-up rule is inconsistent with standard random assignment practice which follows everyone, regardless of their post-randomization decisions. However, the HIP design would yield valid random assignment estimates either (i) if attrition was not related to HIP status or (ii) if differential attrition could be controlled for with sampling weights. Both these assumptions seemed plausible on an a priori basis. The HIP incentive was small relative to the SNAP benefit (and total household income) and therefore unlikely to affect SNAP exit decisions. Furthermore, the assumption was testable. In July 2012 (the end of the Round 2 interview period), the two groups sampled for the survey differ by 5 people (from a total in each group of slightly more than 2,000; this represents one-quarter of a percent). In other months, the differences never rise above 1.5% (slightly more people in the HIP than in the non-HIP group). At randomization, the two non-survey groups were of different sizes, but again the ratio of HIP to non-HIP is stable (at about 10.9%). We conclude that the small observed differences are consistent
with the assumption that differential attrition will not affect the validity of the research design (see Bartlett, Klerman et al. 2013). Finally, the survey data were reweighted to eliminate any imbalance in observable characteristics that might have resulted from differential attrition—within or between the HIP and non-HIP groups—over time.

Data Collection

Dietary recall data were not collected at baseline (Round 1), but were collected in the two post-implementation surveys (Round 2 and Round 3; 4-6 and 9-11 months after implementation, respectively). To support analyses of usual intake (vs. reported intake in the last 24 hours), a 10% subsample received a second interview at each follow-up (Dodd et al. 2006).

This paper analyzes interim dietary intake results from the first follow-up interview (Round 2). That survey sampled 2,784 households, approximately equally split between HIP and non-HIP groups (1,388 and 1,396, respectively). From that sample, 1,998 households completed both survey rounds; again, approximately equally split between treatment and control (1,004 and 994, respectively), with 1,871 respondents providing valid dietary recall data. The final number of interviews in this analysis sample is slightly larger due to the inclusion of the 10% usual-intake subsample (N=2,081).

Measuring TFV Intake

Outcome data were recorded using USDA’s Automated Multiple Pass Method (AMPM) dietary recall system. AMPM is the industry standard for 24-hour dietary recall, and is the method used in the National Health Nutrition and Examination Survey (NHANES) which is the source for the standard information on dietary intake of Americans.
We report the impact of HIP for various fruit and vegetable aggregates. Figure 1 helps to motivate these aggregate groups. The entire oval represents all fruits and vegetables. The light-colored oval in the lower region represents TFVs, excluding 100% juice, white potatoes, and other fruits and vegetables acquired outside stores.

Certain mixed foods in the dietary recall data presented a classification challenge (illustrated at the right side of Figure 1): some mixed foods are fruits and vegetables that were purchased as separate ingredients (qualifying as TFVs) before being mixed in home cooking and preparation, while other mixed foods were purchased in prepared form (not qualifying as TFVs because, for example, they contain added sugar or other ingredients). Our data do not allow us to distinguish between these two possibilities. This leaves us with two choices for constructing our TFV proxy: we can either include all intake of fruits and vegetables from mixed foods (the entire right-hand circle in Figure 1), or exclude all such intake, since we have no practical way to divide up these foods into TFV and non-TFV. The two resulting measures effectively establish a lower bound and an upper bound for actual TFV intake. The former measure excludes some TFVs in mixed foods, while the latter includes some non-TFVs. Before seeing the results, we specified the lower-bound measure as our confirmatory outcome measure of TFV intake, to ensure our preferred measure captured only foods that actually qualified to earn the HIP incentive. Our secondary analyses additionally include impact estimates for the alternative, upper-bound measure.
Sample Statistics

Table 1 reports weighted sample statistics for the survey sample at the first follow-up (Round 2): overall, for the HIP group (i.e., the treatment group in the survey sample), and for the non-HIP group (i.e., the control group in the survey sample). As expected given random assignment and minimal differential attrition, the two groups are balanced on covariates.

Estimates

We estimate the impact of HIP via weighted linear regression using Stata. Covariates include all variables used in blocking and stratification (geography within Hampden County, household composition, and gender of household head), as well as respondent demographic characteristics (age group, gender, race/ethnicity), measures related to baseline consumption according to the Eating at America’s Table Fruit and Vegetable Screener (EATS), covariates about the AMPM interview itself, including day of interview (first or second), and the respondent’s assessment of the prior day’s consumption relative to usual levels (more, less, or the same as usual). All estimates are regression adjusted, using the weighted mean of covariates, as measured at baseline (where the mean is computed pooling across the treatment group and the control group). We compute robust standard errors to account for any heteroskedasticity and the correlation induced by the 10% usual intake sample (i.e., two interviews for the same household).

Figure 2 presents the key impacts on Round 2 daily adult (aged 16 and older) consumption for our preferred TFV intake measure, the evaluation’s pre-specified
confirmatory outcome. TFV intake was slightly more than one-fifth of a cup (0.215 cup-equivalents) higher among HIP participants as compared to those not selected to participate in HIP. This represents a difference of 24.5% over consumption in the absence of the program (0.88 cup-equivalents for control group members). Furthermore, we can decisively reject the null hypothesis of no impact (p=0.001).

Table 2 presents impacts for a wider range of fruit and vegetable aggregates. The difference between our preferred, lower-bound TFV intake measure and our alternative, upper-bound TFV intake measure is substantial. In the control group, regression-adjusted mean intake is 49% higher for our alternative upper-bound TFV intake measure than for our preferred lower-bound TFV intake measure (1.309 cup-equivalents vs. 0.878 cup-equivalents).

Intake of fruits and vegetables from mixed foods, which comprise the difference between our preferred and alternative TFV intake measures, did not significantly differ between HIP participants and non-participants (p=0.848). As a result, impacts on the alternative upper-bound TFV intake measure were similar in magnitude and significance to impacts on our preferred lower-bound TFV intake measure. This result suggests that HIP is primarily influencing intake of TFV consumed alone, without added sugars, fats, oils, or other ingredients.

Considering progressively more inclusive aggregates in turn provides suggestive evidence on which components contribute to HIP impacts on overall fruit and vegetable
intake, beyond TFVs considered alone. As already noted, HIP did not have an impact on fruit and vegetable intake from mixed foods.

Because neither 100% fruit juice nor white potatoes qualify for the HIP incentive, both TFV intake proxy measures exclude these items. However, as seen in Table 2, both 100% fruit juice and white potatoes contribute substantially to total fruit and vegetable intake. Individuals in our control group consumed 0.477 cup-equivalents of 100% fruit juice, comprising about 20% of all fruit and vegetable intake, and 0.351 cup-equivalents of white potatoes, comprising about 15% of all fruit and vegetable intake.

Interestingly, HIP participants consumed significantly more 100% fruit juice than non-participants (0.117 more cup-equivalents), even though fruit juice does not qualify for the HIP incentive. Perhaps some HIP participants thought that 100% fruit juice purchases earned the incentive. Or, perhaps HIP marketing induced a spillover effect by emphasizing the benefits of fruit consumption more broadly, including juice. In contrast, there was no statistically significant impact on white potato consumption.

The TFV intake proxy measures additionally exclude from all fruit and vegetable intake those foods acquired from restaurants, cafeterias, food pantries, or any other source besides a retail store. These foods contribute to the difference in intake between the alternative upper-bound TFV intake measure and all fruit and vegetable consumption not otherwise attributable to intake from white potatoes or 100% fruit juice. This residual comprises about 0.251 cup-equivalents, or about 10% of all fruit and vegetable intake.

Finally, HIP participants’ total fruit and vegetable consumption is approximately one third of a cup (0.343 cup-equivalents) greater than that of non-participants. This impact is somewhat larger than the one fifth of a cup impact on preferred and alternative
TFV intake measures as described above (0.215 and 0.209 cup-equivalents, respectively). The 0.117 cup-equivalents impact on 100% fruit juice intake explains the bulk of the difference in impacts between these two measures. These findings suggest that HIP may have an impact on total fruit and vegetable consumption beyond its direct effects on TFV consumption, though it is unclear why these indirect impacts would be concentrated on 100% fruit juice. There is no evidence of a shift in consumption of fruits and vegetables from non-store sources (which do not qualify for the incentive) to stores (which qualify for the incentive). Specifically, the estimated impact on “other fruits and vegetables acquired outside stores” is small (0.006) and not statistically significant (p=0.837).

We estimated subgroup models by interacting the treatment dummy with several binary subgroup indicators: respondent gender (males vs. females); respondent age group (age 16-40 years vs. 41+ years); primary shopper employment status (employed full- or part-time vs. not employed); SNAP benefit size ($200 or less vs. over $200); and household composition (households with children vs. households with no children). There was no evidence of differential impacts by demographic subgroup.

**Discussion of Interim Pilot Results**

This paper projected the likely impact of HIP given economic theory and the literature on price elasticities. The literature suggests an elasticity of about two-thirds. Therefore, the projection suggested an impact on TFV intake of approximately a fifth of a cup. Several considerations—imperfect understanding and incomplete retailer participation—suggested a smaller estimate. Marketing suggested a larger estimate. Finally, various
forms of uncertainty about the elasticity estimate suggest that these projections are quite uncertain.

In fact, we estimate an impact on TFV intake of 0.215 cups in the Healthy Incentives Pilot. If we were to interpret HIP as a pure price decline, this would imply an elasticity of 0.80, slightly larger than the literature’s estimate of about two-thirds. The estimate is clearly different from zero (p=0.001), but the confidence interval is wide; a 95% confidence interval spans 0.090 cups to 0.340 cups. The corresponding elasticities range from 0.33 to 1.26.

However, as discussed in the second section of the paper, HIP is not a pure price decline, so, interpreting the estimates as an estimate of the elasticity seems inappropriate. In that sense, these results speak to Heckman’s argument in his Nobel lecture (Heckman 2001) for evaluations that estimate behavioral parameters and using those behavioral parameters to forecast the impact of arbitrary policies. In broad form, HIP is a price decline and thus seems to be a natural fit for the analysis that Heckman advocates. Dong and Lin (2009) use exactly that approach to forecast the impact of a HIP-like price reduction, as does the first part of the second section of this paper.

However, the analysis in this paper suggests that Heckman’s recommendations may be, for two reasons, difficult to implement. First, while HIP seems close to a pure price change, the balance of the second section emphasizes just how far from a pure price incentive HIP is. Thus, even if we knew the value of the appropriate elasticity exactly, that might not be sufficient to estimate the impact of HIP.

Second and conversely, even if HIP was a pure price incentive, the HIP evaluation provides little information about the value of the elasticity. The experiment
was deliberately sized to detect “an impact”; it has done so. Precisely estimating the impact and the corresponding elasticity would require a sample much, much larger. The HIP evaluation yields an elasticity with a 95% confidence interval of 0.93 (0.33-1.26). One might plausibly argue that the evaluation would contribute meaningfully to our knowledge of the price elasticity of fruit and vegetable consumption if the confidence interval was 0.20. Cutting the confidence interval by a factor of 4.65 would require a sample more than 20 times as larger (precision declines with the square of the sample size $4.65^2 = 21$). Data collection costs drive evaluation costs, so such an evaluation would be several times more expensive than the one that was funded.

In summary, these estimates suggest that a HIP-like program is a promising strategy for moderately increasing fruit and vegetable intake. The estimated approximately one-fifth cup increase is consistent with other interventions. To put this result into context, consider that the federal government’s Healthy People 2020 objectives (US DHHS 2010) recommend total daily fruit and vegetable (not only TFVs) intake of 3.68 cup-equivalents, which compares to mean estimated intakes of 2.39 cup-equivalents in our non-HIP group, and 2.73 cup-equivalents in our HIP group. So, the HIP impact of 0.34 cup-equivalents was sufficiently large to narrow the 1.29-cup-equivalent “total fruit and vegetable intake gap” by 26%. If the goal is to bring fruit and vegetable intake up to recommended levels, some other intervention—perhaps including a larger incentive—appears to be needed.

Finally, this paper has only considered the impact on TFV intake through the first follow-up interview, about half-way through the pilot. The evaluation includes a second follow-up interview. Including those interviews will allow more precise estimates and the
examination of maturation of the program over time; i.e., do impacts grow or shrink as households gain experience with the incentive. In addition, the full evaluation includes a process analysis (see Bartlett, Klerman et al. 2013) and a cost analysis (to be included in the final report).
This is a best guess based on the disaggregated CEX categories. It excludes raw potatoes and fruit juices. It is, nevertheless, probably an overestimate. Some of the frozen and canned fruits and vegetables almost certainly include potatoes and TFVs with impermissible additives.

The AMPM dietary recall collects information on where food consumed was acquired. However, the difference between impacts on TFV intake and impacts on total fruit and vegetable intake was significant only at the 90% confidence level; in other words, we cannot reject the hypothesis that impacts on TFV intake and impacts on total fruit and vegetable intake were identical in magnitude at the conventional 95% confidence level.

Healthy People 2020 recommends 2.0 cups of fruits and vegetables per 1,000 kcal of daily food energy. Mean daily food energy intake for the non-HIP group was 1,839 kcal.
References


Cole, N., and M.K. Fox. 2008. *Diet Quality of Americans by Food Stamp Participation Status: Data from the National Health and Nutrition Examination Survey 1999-


Figure 1. Targeted fruits and vegetables (TFV) as a proportion of all fruits and vegetables
Figure 2. Impact of HIP on consumption of targeted fruits and vegetables,\textsuperscript{a} mean cup-equivalents consumed

\textsuperscript{a} Targeted fruits and vegetables includes fruits and vegetables acquired from the store, excluding white potatoes, 100% juice, and mixed foods.

Source: AMPM Dietary Recall Interview (unweighted N=2,080 recalls from 1,870 respondents)
## Table 1. Baseline Characteristics of Households Completing Round 2

**Participant Survey, by Treatment and Control Status**

<table>
<thead>
<tr>
<th></th>
<th>Total (N)</th>
<th>Treatment (N)</th>
<th>Control (N)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DTA SNAP caseload data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household residence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Springfield</td>
<td>0.52 (1030)</td>
<td>0.53 (505)</td>
<td>0.52 (525)</td>
<td>[0.698]</td>
</tr>
<tr>
<td>Chicopee or Holyoke</td>
<td>0.25 (505)</td>
<td>0.26 (255)</td>
<td>0.25 (250)</td>
<td></td>
</tr>
<tr>
<td>Hampden County balance</td>
<td>0.22 (419)</td>
<td>0.21 (220)</td>
<td>0.22 (199)</td>
<td></td>
</tr>
<tr>
<td>Persons in household</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SE)</td>
<td>2.33 (0.04)</td>
<td>2.26 (0.05)</td>
<td>2.35 (0.05)</td>
<td>[0.231]</td>
</tr>
<tr>
<td>Single-member household</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One person in household</td>
<td>0.45 (872)</td>
<td>0.44 (450)</td>
<td>0.45 (422)</td>
<td>[0.822]</td>
</tr>
<tr>
<td>Multiple persons in household</td>
<td>0.55 (1082)</td>
<td>0.56 (530)</td>
<td>0.55 (552)</td>
<td></td>
</tr>
<tr>
<td>Adults in household</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SE)</td>
<td>1.49 (0.02)</td>
<td>1.42 (0.02)</td>
<td>1.50 (0.03)</td>
<td>[0.023]**</td>
</tr>
<tr>
<td>Number of adults</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 or fewer adults in household</td>
<td>0.97 (1918)</td>
<td>0.98 (965)</td>
<td>0.97 (953)</td>
<td>[0.055]*</td>
</tr>
<tr>
<td>4 or more adults in household</td>
<td>0.03 (36)</td>
<td>0.02 (15)</td>
<td>0.03 (21)</td>
<td></td>
</tr>
<tr>
<td>Household composition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elderly in household</td>
<td>0.11 (238)</td>
<td>0.11 (129)</td>
<td>0.11 (109)</td>
<td>[0.967]</td>
</tr>
<tr>
<td>Children in household</td>
<td>0.42 (827)</td>
<td>0.43 (398)</td>
<td>0.42 (429)</td>
<td></td>
</tr>
<tr>
<td>No children/elderly</td>
<td>0.46 (889)</td>
<td>0.46 (453)</td>
<td>0.47 (436)</td>
<td></td>
</tr>
<tr>
<td>Monthly SNAP benefit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$161 or less</td>
<td>0.24 (498)</td>
<td>0.25 (257)</td>
<td>0.24 (241)</td>
<td>[0.944]</td>
</tr>
<tr>
<td>$162−$200</td>
<td>0.31 (592)</td>
<td>0.31 (308)</td>
<td>0.32 (284)</td>
<td></td>
</tr>
<tr>
<td>$201−$349</td>
<td>0.14 (272)</td>
<td>0.14 (134)</td>
<td>0.14 (138)</td>
<td></td>
</tr>
<tr>
<td>$350†</td>
<td>0.31 (592)</td>
<td>0.30 (281)</td>
<td>0.31 (311)</td>
<td></td>
</tr>
</tbody>
</table>

*Participant Survey data (primary shopper module)*
# Household composition

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Treatment</th>
<th>Control</th>
<th>P-value</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proportion</td>
<td>Proportion</td>
<td>Proportion</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N)</td>
<td>(N)</td>
<td>(N)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persons in household</td>
<td>2.87 (0.05)</td>
<td>2.82 (0.05)</td>
<td>2.88 (0.06)</td>
<td>0.698</td>
<td>1849</td>
</tr>
<tr>
<td>Children under age 5</td>
<td>0.28 (0.02)</td>
<td>0.32 (0.02)</td>
<td>0.28 (0.02)</td>
<td>0.076*</td>
<td></td>
</tr>
<tr>
<td>Children age 5-17</td>
<td>0.83 (0.03)</td>
<td>0.80 (0.04)</td>
<td>0.84 (0.04)</td>
<td>0.739</td>
<td></td>
</tr>
<tr>
<td>Adults age 18-64</td>
<td>1.58 (0.03)</td>
<td>1.52 (0.03)</td>
<td>1.59 (0.03)</td>
<td>0.169</td>
<td></td>
</tr>
<tr>
<td>Adults age 65 and up</td>
<td>0.18 (0.01)</td>
<td>0.18 (0.01)</td>
<td>0.18 (0.01)</td>
<td>0.924</td>
<td></td>
</tr>
</tbody>
</table>

# Household employment status

(prior week)

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Treatment</th>
<th>Control</th>
<th>P-value</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proportion</td>
<td>Proportion</td>
<td>Proportion</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N)</td>
<td>(N)</td>
<td>(N)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any members full-time</td>
<td>0.19 (320)</td>
<td>0.17 (150)</td>
<td>0.19 (170)</td>
<td>0.327</td>
<td>1829</td>
</tr>
<tr>
<td>employed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any members part-time</td>
<td>0.13 (255)</td>
<td>0.15 (135)</td>
<td>0.13 (120)</td>
<td>0.229</td>
<td></td>
</tr>
<tr>
<td>employed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any members not</td>
<td>0.67 (1261)</td>
<td>0.68 (645)</td>
<td>0.67 (616)</td>
<td>0.589</td>
<td></td>
</tr>
<tr>
<td>employed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

# Program participation

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Treatment</th>
<th>Control</th>
<th>P-value</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proportion</td>
<td>Proportion</td>
<td>Proportion</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N)</td>
<td>(N)</td>
<td>(N)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Received WIC (prior month)</td>
<td>0.18 (335)</td>
<td>0.20 (169)</td>
<td>0.18 (166)</td>
<td>0.509</td>
<td>1842</td>
</tr>
<tr>
<td>Used food pantry/soup</td>
<td>0.12 (209)</td>
<td>0.10 (101)</td>
<td>0.12 (108)</td>
<td>0.414</td>
<td>1841</td>
</tr>
<tr>
<td>kitchen (prior month)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Received Sr. Nutrition/</td>
<td>0.02 (32)</td>
<td>0.02 (17)</td>
<td>0.02 (15)</td>
<td>0.918</td>
<td>1847</td>
</tr>
<tr>
<td>Wheels (prior month)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child received free/reduced price lunch (prior week)</td>
<td>0.87 (702)</td>
<td>0.90 (349)</td>
<td>0.86 (353)</td>
<td>[.] b</td>
<td>789</td>
</tr>
</tbody>
</table>

Note: Weighted proportions (unweighted Ns) for categorical variables; weighted means (standard errors) for continuous variables. Chi-square test for categorical variables, t-test for continuous variables; two-sided test:

* p<0.1, ** p<0.05, *** p<0.01. Due to rounding, reported proportions may not sum to one.

a Top-coded at 7

b Missing test statistics because of stratum with single sampling unit

c “Don’t know,” “refused,” “inapplicable,” and “not ascertained” responses in Participant Survey (primary shopper module) data coded as missing

d Households with qualifying child only

Sources: DTA SNAP Caseload Data (July 2011), Participant Survey (primary shopper module)
Table 2. Impact of HIP on Consumption of Fruits & Vegetables and Disaggregated Components, Cup-Equivalents

<table>
<thead>
<tr>
<th></th>
<th>Regression-adjusted mean (S.E.)</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment (T)</td>
<td>Control (C)</td>
</tr>
<tr>
<td>Targeted fruits and vegetables (preferred lower-bound proxy)¹</td>
<td>1.093 (0.050)</td>
<td>0.878 (0.041)</td>
</tr>
<tr>
<td>MTFV from mixed foods</td>
<td>0.425 (0.023)</td>
<td>0.431 (0.022)</td>
</tr>
<tr>
<td>Targeted fruits and vegetables (alternative upper-bound proxy)ᵃ</td>
<td>1.518 (0.054)</td>
<td>1.309 (0.048)</td>
</tr>
<tr>
<td>Additional components:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100% fruit juice</td>
<td>0.594 (0.041)</td>
<td>0.477 (0.031)</td>
</tr>
<tr>
<td>White potatoes</td>
<td>0.361 (0.024)</td>
<td>0.351 (0.022)</td>
</tr>
<tr>
<td>Other fruits and vegetables acquired outside stores</td>
<td>0.257 (0.019)</td>
<td>0.251 (0.021)</td>
</tr>
<tr>
<td>All fruits and vegetables</td>
<td>2.731 (0.075)</td>
<td>2.388 (0.069)</td>
</tr>
</tbody>
</table>

Note: Two-sided test; *p<0.1, **p<0.05, ***p<0.01. Due to rounding, reported impacts (T-C differences) may differ from differences between reported regression-adjusted means for the treatment and comparison groups.

ᵃ Targeted fruit and vegetable (TFV) intake proxy measures include intake of fruits and vegetables acquired from the store, excluding white potatoes, legumes, and 100% juice. The preferred lower-bound proxy measure additionally excludes fruit and vegetable intake from mixed foods where the source of individual ingredients was not identified by the respondent, while the alternative upper-bound proxy measure includes fruit and vegetable intake from all mixed foods.

Source: AMPM Dietary Recall Interview (unweighted N=2,080 recalls from 1,870 respondents)