Would a discount on fruits and vegetables provide more relative welfare to the poor? Evaluating the impact of policy mechanisms.

Dominique J Rolando, University of Minnesota (rolan059@umn.edu)

Selected Paper prepared for presentation at the 2018 Agricultural & Applied Economics Association Annual Meeting, Washington, D.C., August 5-August 7

Abstract

Food assistance is a highly controversial topic in the U.S., especially given that big programs such as the Supplemental Nutrition Assistance Program (SNAP) cost billions of Dollars. Though these programs help promote food security, a concern for the provision of healthy diets to program participants remains. This paper compares the simulated impacts of food assistance similar to SNAP to two alternative policy mechanisms: a cashback program whereby participants are reimbursed for a set proportion of their fruits and vegetable expense, and a discount on the purchase price of fruits and vegetables. Using Nielsen Homescan data, a Quadratic Almost Ideal Demand System is estimated to obtain the necessary parameters to obtain the impacts of these policies. The results make it apparent that households at high levels of poverty benefit more from assistance programs similar to SNAP, while those above the poverty line could benefit more from both the cashback and discount. The discount also appears to provide better gains to participating households than the cashback program.

1. Introduction

Food security is usually measured in absolute consumption of food and the ease with which households acquire food. Though this approach is quite reasonable, less regard is given to the composition of one's diet. As such, food assistance has long focused on the provision of enough calories to sustain daily activities rather than on which nutrients were included in the calories. In fact, diets variy largely by income status due to the impact of food prices (Lin, 2005; Drewnowski, and Eichelsdoerfer, 2010), among other economic forces. Additionally, prices and budgeting are viewed as barriers for proper diet by low-income households (Dachner, Ricciuto, Kirkpatrick, & Tarasuk, 2010), who would expect to buy healthier options were they to have a larger food budget (Inglis, Ball, & Crawford, 2009).

A food group that is under-consumed in the American diet is fruits and vegetables. The Dietary Guidelines Advisory Committee, a joint effort between the United States Department of Agriculture and the United States Department of Health and Human Services, issued a report in 2015 outlining the under-consumption of fruits, vegetables, and wholegrains by Americans. In the same report, they found that sugars and fats were overconsumed, contributing to the rise of chronic diet related diseases. When analyzing energy density of foods and income levels, Drewnowski and Specter (2004) found that the overconsumption of energy dense foods (sugars, refined grains) were relatively high at low incomes, which they attributed to prices of those goods being within reach for the poor. Conversely, fruits and vegetables were under-consumed. It is of no help that the worse foods for one's diets are also more palatable for many.

It is thus clear that those of low-income, like many Americans, do not consume enough nutritious foods but consume too much sugar and fats. Though food assistance programs, such as the Supplemental Nutrition Assistance Program (SNAP), provide funds for low income Americans to afford food, the income increase might not prove enough to overcome the relative cost of healthy diets. The question that I pose then is: How would low-income households gain from other food assistance mechanisms that promote healthy food purchases?

The main objective of this paper is to compare the impact of three policy mechanisms on low-income households: needs-based food assistance (similar to SNAP), a cashback program, and a price discount. Using Nielsen Homescan data for the year 2016, I aggregate foods into intuitive groups (fruits and vegetables, meats and dairy, fats, sugar and confectionery goods, all others). Fruits and vegetables are chosen as the "token" healthy food group since their healthfulness is intuitive, and they are under-consumed by U.S. households, as mentioned above. I then simulate the impact of the aforementioned policy mechanisms on each household and compare how they affect said households. In order to obtain the impact of the price discount, I estimate consumer demand using a Quadratic Almost Ideal Demand System, then simulate a price discount on fruits and vegetables. Using model parameters, I calculate the associated Compensating Variation for each household. The hypothesis is that the most effective policy mechanism will be different at various points along the poverty threshold. The contribution of this paper lies in the illustration of how policy aimed at encouraging healthy purchases affect low-income households relative to needs-based food assistance.

The paper progresses as follows: Section 2 provides helpful background information surrounding the topics covered in the paper; Section 3 defines the conceptual framework behind the methodology used and outlines the estimation strategy; Section 4 covers the data, illustrates general data composition, and the procedure involved in constructing group prices; Section 5 covers results; Section 6 discusses policy implications; and Section 7 provides concluding remarks.

2. Background

2.1. Food, poverty, and dieting.

Semega, Fontenot, & Kollar (2017) report to the United States Census Bureau that the poverty rate in the U.S. for the year 2016 was 12.7%. Poverty is linked to multiple social problems, of which a key issue is that of food security. The United States Department of Agriculture (2017) described very low food security as “reports of multiple indications of disrupted eating patterns and reduced food intake.” The striking part of this definition is that it centers on food intake with no mention of the type of food being ingested. A household could thus be food secure but be nutritionally poor. To further display this point, studies have found that low-income households had less nutritious diets than higher income households (Jones, Akbay, Roe, & Chern, 2003) and were least likely to purchase foods that had high fiber, and low fat, sugar, and salt (Turrell and Kavanagh, 2006). Even though the overall healthfulness of diets of most households in the U.S. are improving (Beatty, Lin, & Smith, 2014), the quality gap between rich and poor households is increasing (Wang et al., 2014). Nutrition is thus of importance to low-income households as poor nutrition is linked to numerous chronic diseases. It is estimated that, in the year 2000, about 17% of deaths in the U.S. were caused by poor diet and physical activity (Mokdad, Marks, Stroup, & Gerberding, 2004).

Market forces appear to obstruct low income households attempting to obtain a nutritious diet. Not only is their income limited, but the sheer cost of a nutritious diet is higher in food deserts (Fan, Baylis, Gundersen, & Ver Ploeg, 2015, which are census tract areas with more than 20% poverty rate and where a significant portion of residents live far from a food retail location (United States Department of Agriculture, 2017). To assist with this issue, food assistance programs such as SNAP provide funds to qualifying low income households to purchase food. A common belief is that further restricting SNAP eligible foods to stop unhealthy food purchases might help improve SNAP participants' diets. Such an approach would likely reduce the effectiveness of the program at countering food insecurity (Gregory, Ver Ploeg, Andrews, & Coleman-Jensen, 2012).

Therefore, public policy should consider alternative mechanisms and programs if the goal is to promote healthful diets. The ideal scenario is to use a policy mechanism that has an impact equal to at least equal that of a program such as SNAP. An alternative to food budget assistance is that of subsidizing the price of certain healthy foods, effectively reducing their price. Dong and Lin in their 2009 report to the United States Department of Agriculture found that a 10% subsidy on the prices of fruits and vegetables could potentially increase their intake by low income households by 2.1-5.2 percent. Another alternative is a cashback program where participants are awarded a set percentage of what they spend on fruits and vegetables as additional money to be spent on food, similarly to the United States Department of Agriculture's Healthy Incentive Program (HIP). HIP was found to have a positive impact on the purchase of fruits and vegetables for participants (Bartlett., & Abt Associates, 2014).

Though these programs were evaluated on how effective they are at promoting healthy food purchases, little regard has been given to how they each impact the economic welfare of participating households. The purpose of this paper is to evaluate the impact of each of these programs on low-income households, comparing their impact, and evaluating whether any one of them helps households more than the other.

2.2. Income and consumption.

Since the days of Engel (1895), and later to Working (1943) and Leser (1963), the relationship between expenditure and income have been central to the analysis of consumer behavior. According to Engel's law, as income increases the budget share spent on food decreases, even if the nominal amount of income increases. In this regard, the complex relationships between changes in income and food expenditure share have been of great interest to economists. This body of work has looked at various aspects of consumer spending with early measurement efforts by Davidson, Hendry, Srba, & Yeo (1978) and has ranged from alcohol consumption (Atkinson, Gomulka, & Stern, 1990), to more recent efforts on healthcare expenditure in OECD countries (Baltagi and Moscone, 2010).

Often to analyze expenditure changes at various income levels Engel curves are used. Engel curves display the relationship between household expenditure on a particular good at varying levels of income. Burzig, and Herrmann (2012) used the Engel curve to analyze spending patterns on food-at-home and food-away-from-home in Germany. They found that the theoretical expectations hold true and budgets behave as expected with changes in income. That is, as income increases, so does spending on food-at-home. However, the share spent on food-at-home goes down with additional income. Using a comparable approach, Magana-Lemus and co-authors (2013) found that an increase in tortilla prices affected low income Mexican households almost twice as much as higher income households. Gale and Huang (2007) found that the nominal amount spent on food in China is growing faster than the quantity of food bought. They attributed this phenomenon to Chinese consumers opting for higher quality of food over quantity.

In order to better understand the relationship between consumer expenditure on specific goods and income, it is often necessary to estimate consumer demand. Deaton and Muellbauer's (1980) Almost Ideal Demand system is a popular approach to demand estimation but their specification is not flexible enough to accommodate non-linear Engel relationships. Complex Engel relationships were then operationalized by Banks, Blundell, and Lewbell (1997) using a Quadratic Almost Ideal Demand System. They found that Engel relationships were in fact non-linear for some commodity groups, notably clothing and alcohol. In intuitive terms, the income elasticity of clothing revealed it was a luxury for some income levels but a necessity for others. This prompted a need for the use of higher order terms in demand estimations to allow for flexible Engel curves. The innovation of a flexible demand specification proved relevant to numerous consumer choice analyses, notably on revealed preferences (Blundell, Browning, & Crawford, 2003), welfare evaluation (Banks, Blundell, & Lewbel, 1996), household allocations and bargaining (Browning and Chiappori, 1998; Browning, Chiappori and Lewbel, 2006), as well as investigations into shape invariant Engel curves (Blundell, Chen, and Kristensen, 2007).

3. Conceptual framework and estimation models

3.1. Empirical strategy

The main goal of this paper is to compare the impact of policy mechanisms aimed at encouraging the purchase of fruits and vegetables, an under-consumer healthy food. The two policy mechanisms considered are a cashback program on fruits and vegetable purchases and a discount on the price of fruits and vegetables. Before evaluating these mechanisms, it is important to establish a baseline. The appropriate baseline in this case is a representation of the status-quo. Currently SNAP is the largest food assistance program in the U.S. SNAP provides funds to participants which they can use to purchase food. The amount of funds they receive is dependent on their needs: a household below the poverty line receives more assistance than a household above it. It is therefore important that this paper's empirical strategy simulates such a program for comparison sake.

In order to obtain the impact of a needs based food assistance program, a cashback program, and a price discount, one must use a stable unit of measurement. The most straightforward unit of measure in the case of this paper is the associated gains in food expenditure from each of the proposed mechanisms.

In simulating a believable policy, I first start with a budget of $1 billion of fictitious funds available to fund each program. Though this amount is chosen arbitrarily[1], it works well for the purposes of this paper firstly for its intuitive appeal, but also for the fact that it makes up

[1] Changing the fictitious budget would only mean scaling the findings up or down, which should not have much influence on their reliability.

only a small percent of the current federal budget on major food assistance programs such as SNAP. Changing this amount should not affect the reliability of this study's findings drastically.

The next step in the analysis is thus to use the available budget to fund each of the possible policy programs and measuring how households gain from them in terms of food expenditure. Obtaining the gains from needs-based food assistance and the cashback program are quite straightforward. For needs-based food assistance, the further below the poverty line a household is, the more food assistance it receives. For instance, if a household at the poverty line (100% of poverty) receives $100, a household at 50% of poverty would receive $150. Households above 150% of poverty do not receive assistance, consistently with the majority of households eligible for SNAP. For the cashback program, the amount gained by households is a fixed proportion of what they have spent on fruits and vegetables. Using the fictitious budget, the program could fund a cashback of approximately 5% of fruits and vegetable purchases.

Obtaining the impact of the discount is the tricky part. Based on the fictitious budget, a discount of approximately 5% would be feasible. To find the impact of the discount, I estimate consumer demand and use the demand system to obtain the Compensating Variation (CV) associated with the price reduction. This CV is analogous to the impact of the price discount for participating households.

3.2. The Almost Ideal Demand System

In order to properly compute the CV associated with an 5% discount, a demand system must first be estimated. The selection of a functional form is fundamental to estimating demand. Such a functional form must respect the theoretical properties of consumer demand. Deaton and Muellbauer's (1980) Almost Ideal Demand System (AI) has thus become one of the most widely used model for estimating consumer demand.

This system starts with the dual problem of utility maximization, cost minimization. The end result is a demand system that can be estimated and where constraints can be added to respect economic theory. Equation (1 shows the final form of the AI model.

(1) $$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} p_j + \beta_i \log \frac{W}{a(\boldsymbol{p})}$$

where w_i is the budget share of good i and $a(\boldsymbol{p})$ is defined as:

(2) $$a(\boldsymbol{p}) = \alpha_0 + \sum_{k=1}^{n} \alpha_k \log p_k + \frac{1}{2} \sum_{k=1}^{n} \sum_{l=1}^{n} \gamma_{lk} \log p_k \log p_l$$

Price and income elasticities can be computed post estimation by taking the derivative of the log share equation with respect to the log price or log income. Furthermore, it is possible to simulate the welfare change associated with a change in prices. This is done using the CV measure given by:

(3) $$CV(\boldsymbol{p^0}, \boldsymbol{p^1}, W) = C(\boldsymbol{p^0}, u^0) - C(\boldsymbol{p^1}, u^0)$$

where the original prices are $\boldsymbol{p^0}$, the new prices are $\boldsymbol{p^1}$, and *C(.)* is the cost function associated with a given utility and price level. While the term $C(\boldsymbol{p^0}, u^0)$ is simply expenditure, the problematic term in this is $C(\boldsymbol{p^1}, u^0)$. To obtain it, consider the cost version of the AI model:

(4) $$\log C\,(\boldsymbol{p}, u) = a(\boldsymbol{p}) + b(\boldsymbol{p}) * u$$

where $a(\boldsymbol{p})$ is defined in equation (2 and $b(\boldsymbol{p})$ is defined as:

(5) $$b(\boldsymbol{p}) = \beta_0 \prod_{k=1}^{n} p_k^{\beta_k}$$

For original prices and utility equation (4 can be rearranged to obtain the indirect utility function:

(6) $$\frac{(\log C(\boldsymbol{p^0}, u^0) - a(\boldsymbol{p^0}))}{b(\boldsymbol{p^0})} = u^0$$

Similarly,

(7) $$\log C\,(\boldsymbol{p^1}, u^0) = a(\boldsymbol{p^1}) + b(\boldsymbol{p^1}) * u^0$$

Plugging in u^0 from equation (6:

(8) $\log C\,(\boldsymbol{p^1}, u^0) = a(\boldsymbol{p^1}) + b(\boldsymbol{p^1}) * [\frac{(\log C(\boldsymbol{p^0}, u^0) - a(\boldsymbol{p^0}))}{b(\boldsymbol{p^0})}]$

Using the fact that $\frac{b(\boldsymbol{p^1})}{b(\boldsymbol{p^0})} = \frac{\beta_0 \prod_k (p_k^1)^{\beta_k}}{\beta_0 \prod_k (p_k^0)^{\beta_k}} = \left(\frac{p_j^1}{p_j^0}\right)^{\beta_j}$ where j is the good with the price change, the result is:

(9) $C(\boldsymbol{p^1}, u^0) = e^{\{a(\boldsymbol{p^1}) + \left(\frac{p_j^1}{p_j^0}\right)^{\beta_j} * [(\log C(\boldsymbol{p^0}, u^0) - a(\boldsymbol{p^0}))]\}}$

It is thus possible to obtain a measure for CV with associated with a price change.

3.3.The Quadratic Almost Ideal Demand System

The Engel relationship for different goods need to be more flexible than what is captured in the AI model. Banks, Blundell, and Lewbel (1997) recognized this issue and displayed how Engel curves need not be linear at all expenditure levels. In order to accommodate such relationships, the Quadratic Almost Ideal Demand System (QUAI) includes an extra quadratic term in the AI model which allows for more flexibility. The model is similar to the AI model but with an indirect utility function defined as:

(10) $log\ V(\boldsymbol{p}, W) = \left\{\left[\frac{log\ W - log\ a(p)}{b(\boldsymbol{p})}\right]^{-1} + \lambda(\boldsymbol{p})\right\}^{-1},\ where\ \lambda(\boldsymbol{p}) = \sum_{i=1}^{n} \lambda_i \ln\ p_i$

The QUAI model then takes the form:

(11) $w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} p_j + \beta_i \log \frac{W}{a(\boldsymbol{p})} + \frac{\lambda_i}{b(\boldsymbol{p})} \left[\frac{\log W}{a(\boldsymbol{p})}\right]^2$

The last term in equation (11 provides flexibility to this functional form. Similarly to the AI model, elasticities for both price and income can be derived from the QUAI model. The derivation of CV is also slightly different. Though CV is defined exactly the same, the derivation of $C(\boldsymbol{p^1}, u^0)$ is different. From equation (10, the indirect utility function takes a different form than in the AI model. Therefore, the cost function takes the form:

(12) $log\ C(\boldsymbol{p}, u) = log\ a(\boldsymbol{p}) + \frac{b(\boldsymbol{p}) log\ u}{1 - \lambda(\boldsymbol{p}) log\ u}$

Once again plugging the indirect utility function with the original price into $C(\boldsymbol{p^1}, u^0)$ used to calculate CV:

$$(13) \quad log\ C(\boldsymbol{p^1}, u^0) = \log a(\boldsymbol{p^1}) + \frac{b(\boldsymbol{p^1})\left\{\left[\frac{log\ W - \log\ a(\boldsymbol{p^0})}{b(\boldsymbol{p^0})}\right]^{-1} + \lambda(\boldsymbol{p^0})\right\}^{-1}}{1 - \lambda(\boldsymbol{p^1})\left\{\left[\frac{log\ W - log\ a(\boldsymbol{p^0})}{b(\boldsymbol{p^0})}\right]^{-1} + \lambda(\boldsymbol{p^0})\right\}^{-1}}$$

Taking the exponent of the right-hand side of equation (13, gives $C(\boldsymbol{p^1}, u^0)$ needed to calculate the CV associated with a price change

3.4. Non-linear estimation

The QUAI model needs to be estimated non-linearly because of the non-linearity of the terms in the functional forms. Blundell and Robin (1999) propose a method to estimate similar conditionally linear system using iterated linear least squares.

For the AI model, the problematic term is log $\frac{W}{a(\boldsymbol{p})}$. The AI model could be made linear by approximating *a(**p**)* to the Stone Price Index. However, I use the iterated linear least squares method to approximate *a(**p**)*. In the first iteration, *a(**p**)* is approximated by the Stone Price Index, while α_0 is assigned the minimum of the log expenditure. The system is estimated via Seemingly Unrelated Regressions (SUR) and the parameters that make up *a(**p**)* are replaced by those obtained in the estimation. The system is re-estimated with the new *a(**p**)*. This continues until the system parameters converge. Similarly, for the QUAI model the Stone Price Index is the starting value of the iteration for *a(**p**)* while the starting value for *b*(***p***) is 1. The theoretical requirements of symmetry and homogeneity are imposed in the model via constrained SUR estimation.

The allure of this approach as opposed to simply estimating AI or QUAI non-linearly is the computation of the *a(**p**)* and *b*(***p***) terms in each iteration, which is quite useful when calculating CV.

4. Data

To meet the model requirements described earlier, the data need to include information on consumer purchases, including prices and quantity, as well as consumer income. For a better model, household demographics should also be included as controls, as well as measures of time and location to account for seasonality and geographical differences. The dataset that I use is the Nielsen Homescan Database for the year 2016 as it incorporates all the necessary information.

The data are collected via a shopper panel who record all their purchases and prices. The data also include demographics of the household. Nielsen selects a subset of those who sign up for the program based on the needs of the data. The main goal is to obtain a representative panel of consumers. Participants are then given a scanner with which they scan the barcode of the products they are purchasing. The data are then compiled and distributed by Nielsen.

The strength of the dataset lies in its completeness. Not only does it include helpful household level characteristics, it also includes complete purchase history of participants. Nielsen data are not without flaws however. There are concerns about accuracy, sample selection, as well as misreporting (Einav, Leibtag, & Nevo, 2008; 2010) in the data. The main issue arises from participants misreporting their purchases or from price imputations done by Nielsen, as well as underrepresentation of some income groups. Though these issues are present, there are no clear ways to correct for them as they are structurally present in the data compilation.

4.1. Descriptive statistics

The data include numerous household level demographic details, though not all of them are useful for the purposes of this paper. Table 1 shows a breakdown of demographics used in the AI and QUAI models. Most of the sample is white (~74%) with most female or male heads being above 55 years old. An interesting figure is that over 1/3 of households (~45%) do not have a male head present. The mean household size is slightly above 2, which coincides with most of the households sample having no children present (~66%).

Table 1. Descriptive Statistics of Nielsen Homescan Data 2016.

Descriptive Statistics	Column % or Mean
Non-White	25.6
White	74.4
Female Head	
Not present in hh	20.7
Under 25 Years	1.8
25-29 Years	5.4
30-34 Years	9.8
35-39 Years	6.8
40-44 Years	6.3
45-49 Years	6.7
50-54 Years	8.6
55-64 Years	16.2
65+ Years	17.7
Male Head	
Not present in hh	44.8
Under 25 Years	0.7
25-29 Years	2.9
30-34 Years	6.4
35-39 Years	5.4
40-44 Years	4.9
45-49 Years	4.8
50-54 Years	6.4
55-64 Years	11.8
65+ Years	12.0
No children in hh	65.8
Child present in hh	34.2

Descriptive Statistics	Column % or Mean
Mean hh size	2.6
N	15,139

Note: All figures shown are column percentages, unless specified as a mean. Nielsen Homescan data weights used.
Source: Author's calculations using Nielsen Homescan Data 2016.

4.2. Price construction

Prices are an important component of any demand estimation. The well acknowledged problem with demand estimation is that of dimensionality. To be accurate, demand needs to be estimate for all possible commodities. This is however infeasible. Therefore, to estimate demand, I assume weak separability and treat foods as being part of the first stage of a two-stage budget. The second stage then include food groups that I define as fruits and vegetables, meats and dairy, fats, soda and confectionery goods, and all other goods. This solves the dimensionality problem and allows for estimation. These food groups are due to their intuitive appeal, with fruits and vegetable being the main group of interest for the analysis.

A problem subsides in that each of those groups has many individual products present in them. Though their quantities can be aggregated, coming up with common prices for the group can be complicated. To solve this issue, I calculate the geometric mean of prices for each commodity in the group and assign it as the group price. An issue with using the geometric mean of prices is that I run the risk of having too little price variation, thus coarsening the measure of price levels too much, which dilutes potential effects of geography and seasonality. Therefore, the geometric mean must incorporate a measure of time and geography. I thus calculate the geometric mean at the region and quarter level. In other words, each food group for a region *r* and quarter t will be assigned the geometric mean for all individual commodities in that group for that region and quarter.

The price is calculated as follows:

(14) $$P_{xtr} = e\left(\sum_{j=1}^{n} \frac{\bar{w}_{jtr}}{\sum_{k=1}^{n} \bar{w}_{ktr}} * \log P_{jtr}\right)$$

where P_{xh} is the price for product group x faced by region r at quarter t, and $\bar{w}_{jt}$ is the average budget share of product j (part of group x) at quarter t and in region r. Therefore, the price for each group is unique for a specific region and a specific quarter. A breakdown of prices by region and quarter are shown in Table 2.

Table 2. Prices for each food group by region and quarter.

	Q1	Q2	Q3	Q4
Fruits and vegetables				
New England	2.35	2.37	2.44	2.33
Middle Atlantic	2.46	2.48	2.49	2.47
East North Central	2.16	2.17	2.18	2.15
West North Central	2.31	2.32	2.31	2.27
South Atlantic	2.28	2.32	2.29	2.26
East South Central	2.15	2.15	2.16	2.06
West South Central	2.12	2.13	2.12	2.09
Mountain	2.21	2.22	2.19	2.18
Pacific	2.35	2.36	2.34	2.29
Total	2.26	2.28	2.27	2.24
Meats and dairy				
New England	4.99	5.13	5.15	5.09
Middle Atlantic	4.98	5.07	5.03	5.10
East North Central	4.56	4.68	4.57	4.62
West North Central	4.49	4.51	4.45	4.58
South Atlantic	5.22	5.22	5.18	5.25
East South Central	4.71	4.83	4.71	4.76
West South Central	4.94	5.00	4.94	4.98
Mountain	5.08	5.09	5.12	5.20
Pacific	5.64	5.71	5.67	5.77
Total	5.01	5.08	5.02	5.09

Fats				
New England	3.33	3.32	3.39	3.32
Middle Atlantic	3.36	3.37	3.39	3.38
East North Central	3.13	3.05	3.11	3.09
West North Central	3.25	3.19	3.24	3.17
South Atlantic	3.44	3.35	3.38	3.38
East South Central	3.24	3.16	3.17	3.19
West South Central	3.42	3.39	3.40	3.43
Mountain	3.52	3.48	3.52	3.54
Pacific	3.74	3.74	3.75	3.82
Total	3.40	3.35	3.38	3.39
Soda and confectionery goods				
New England	2.80	2.76	2.73	2.86
Middle Atlantic	2.83	2.79	2.82	2.90
East North Central	2.68	2.65	2.65	2.72
West North Central	2.77	2.75	2.76	2.80
South Atlantic	2.79	2.76	2.75	2.83
East South Central	2.60	2.59	2.57	2.62
West South Central	2.78	2.76	2.74	2.78
Mountain	2.95	2.94	2.91	2.96
Pacific	3.16	3.13	3.13	3.22
Total	2.83	2.80	2.79	2.86

Note: Nielsen Homescan data weights used.
Source: Author's calculations using Nielsen Homescan Data 2016.

4.3. Income

A possible pitfall of Nielsen Homescan Data is the reported income of households. The income figure used is not only presented as a range of possible income, it also is lagged by two years. For the 2016 data, the income available are household incomes for the year 2014. In order to obtain a usable measure of income, I assign each household's income to be a random number in the range of income provided. For example, if a household reports that its income is between $30,000 and $34,999, I assign the household income to be a random number within that range, say $31452. There is however no possible way to correct for the lagged nature of the income ranges reported.

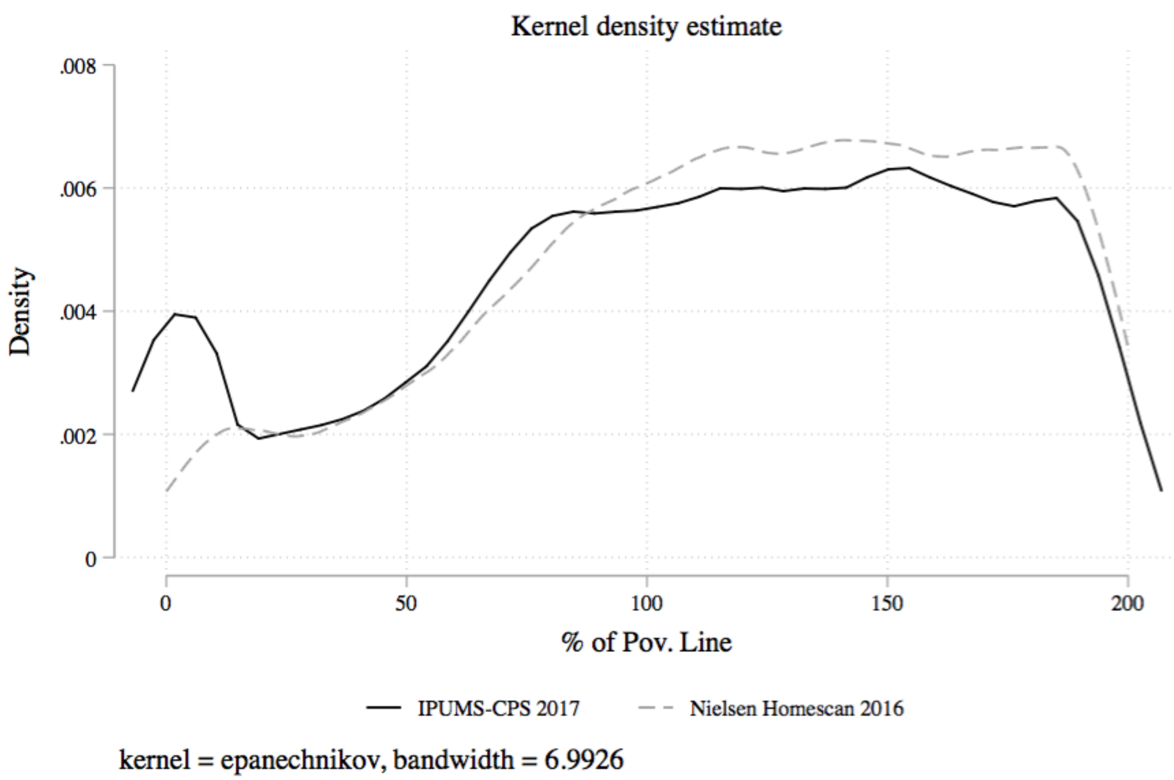
Estimating QUAI requires expenditures. Fortunately expenditures reported in Nielsen Homescan Data 2016 are current to 2016, therefore the lagged income measure should not influence the estimation. The reported income comes into play when calculating the level of poverty of a household. The Census Bureau uses the Poverty Thresholds (United States Census Bureau, 2016) to determine the poverty line in the U.S. The thresholds vary by household size, the number of children, and presence of individuals over the age of 65, which allows for a more accurate depiction of poverty based on household characteristics. Dividing household income by its corresponding poverty thresholds provides a distance metric of how far the household is from the poverty line, either above or below. For instance, the poverty threshold for a household of two without children and no adults over 65 in the year 2016 was $16,072. A household making exactly that amount is at 100% of poverty, one making half of is at 50% of poverty, while one making twice that is at 200% of poverty. For this analysis, only households at 200% of poverty or below are considered.

The poverty threshold measure is important in determining the impact of the proposed different policy mechanism at various levels of poverty. Furthermore, the measure is the sole determinant of the benefit amount given out through the fictitious needs-based food assistance program. That is, a household at 50% of poverty would be getting twice as much in terms of benefits than a household at 100% of poverty. Therefore, in order to calculate it a proper measure of income is needed.

A possible assumption in the case of this study, given the lagged income reporting in the data, is that household income for the households in the sample are representative of households in the year 2016. In order to assert whether this assumption is reasonable, I compare the kernel density of households in the sample to the Current Population Survey (CPS) sample for the year 2016. The goal is to assert whether the Nielsen Homescan 2016 lagged income (therefore income represented in Nielsen is from 2014) as a proportion of poverty properly represents this measure for the U.S. in 2016.

Figure 1 shows the kernel density graph of income calculated as a percentage of the poverty threshold, comparing the Nielsen Homescan 2016 data (income from 2014) and the IPUMS-Current Population Survey data 2017 (income from 2016).

Figure 1. Kernel density comparison graph of Income as a % of the Poverty Threshold for Nielsen and IPUMS-CPS.



Note: Nielsen homescan data weights used; IPUMS-CPS weights used.

Source: Author's calculations using Nielsen Homescan Data 2016, United States Census Bureau Poverty Thresholds 2014/2016, and IPUMS-CPS 2017.

Reassuringly, the kernel density for both datasets appears to track fairly close together. The CPS has a higher density of households at the left tail of the distribution but the density tracks almost perfectly up to 50% of poverty. Then the CPS has a slight higher density up until close to 100% of poverty where Nielsen has a higher density for the rest of the sample. Though it is clear that the Nielsen data does not perfectly represent the poverty distribution of the CPS, it is fairly close. This makes the assumption that the lagged income variable reported in Nielsen for the year 2016 is fairly representative of incomes in 2016.

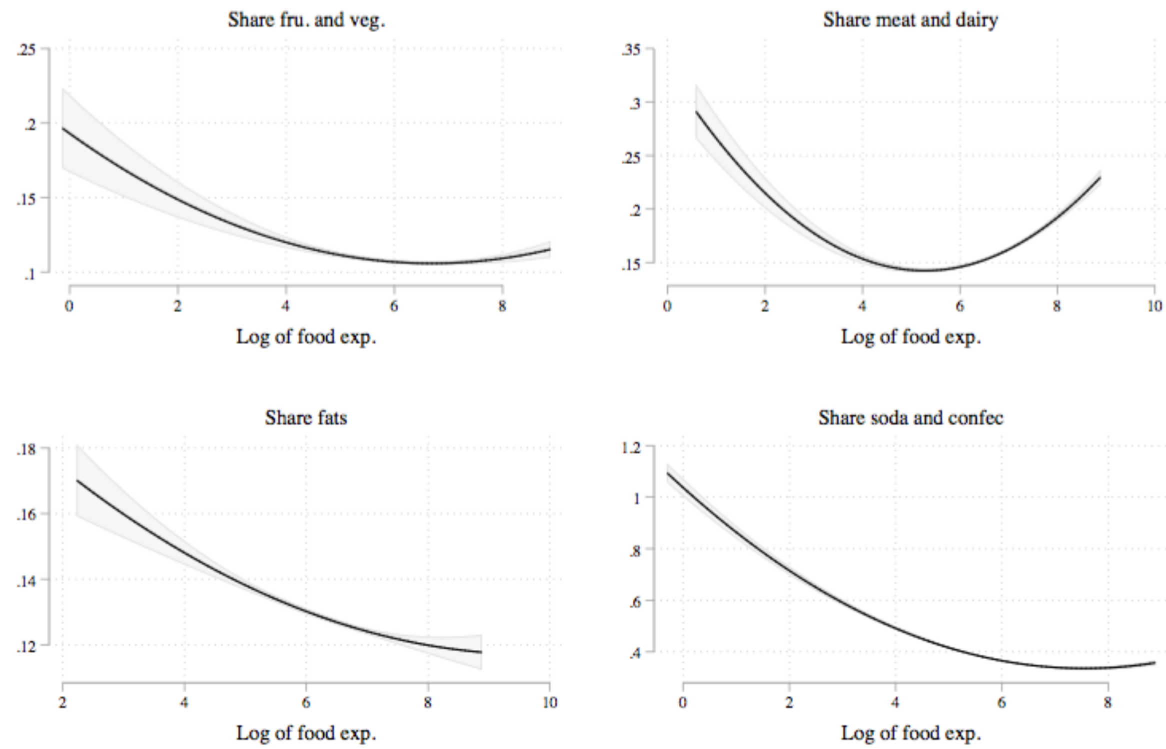
5. Results

Before delving into the impact of policy mechanisms proposed, it is important to consider the various aspects of consumer choice involved which will influence the impact of policies.

5.1. Food expenditure

To obtain a simplistic glimpse at the Engel curves for various food groups and food expenditure, I plot the log of food expenditure against the budget share spent on each food group using a quadratic prediction plot.

Figure 2. Quadratic prediction plots of log food expenditure versus share of budget spent on food group.



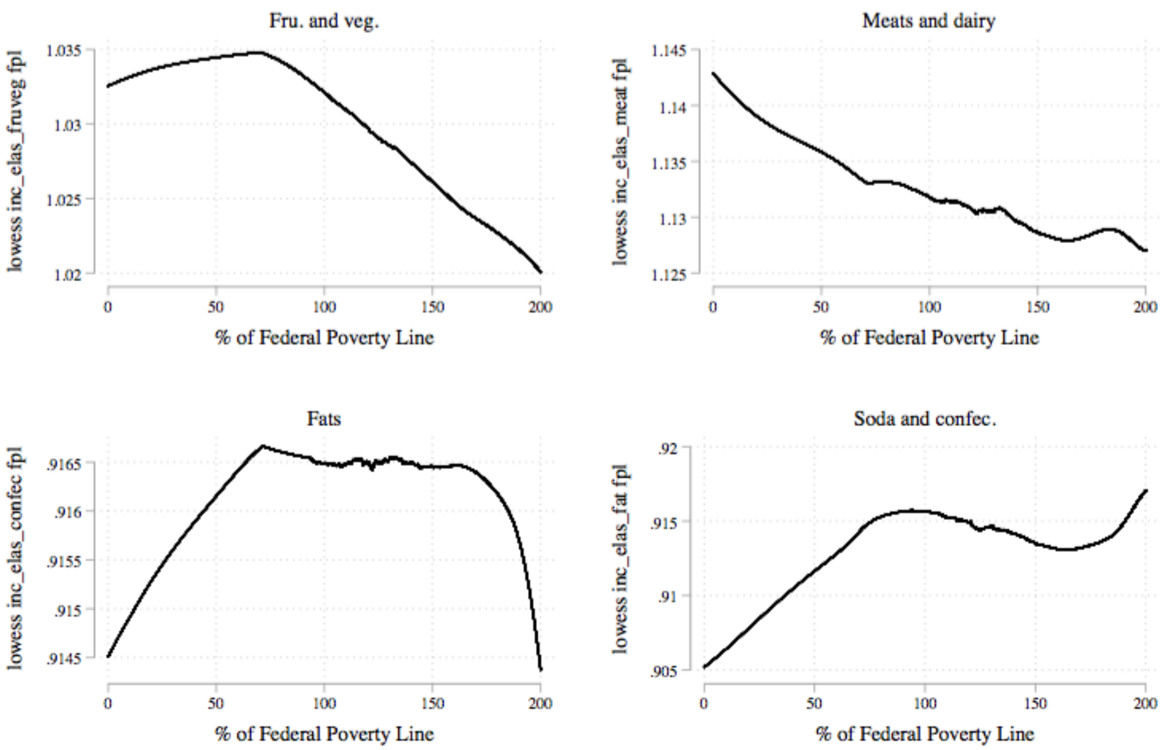
Note: Shaded area represents 95% confidence intervals; Nielsen Homescan data weights used.
Source: Author's calculation using Nielsen Homescan data 2016.

As the log food expenditure increases, the share spent on fruits and vegetables decreases first, with a turning point happening the right-hand side of the distribution indicating a potential rise in fruits and vegetable expenditure. For meats and dairy the turning point happens at lower levels of food expenditure and is more drastic. More general downward trends are observed for fats, and soda and confectionery goods. These graphs show that there is an important aspect of how much of each food group a household purchases being dictated by the expenditure level. More importantly, it is apparent that not all of those Engel relationships are linear, similar to Banks et al.'s (1997) findings.

5.2. Income elasticity

After estimating the QUAI model, I compute income elasticities using model parameters. Figure 3 shows Locally Weighted Scatterplot Smoothing (LOWESS) plots of each food group elasticity against the percentage of poverty line for each household.

Figure 3. Income elasticity of food groups- AI model



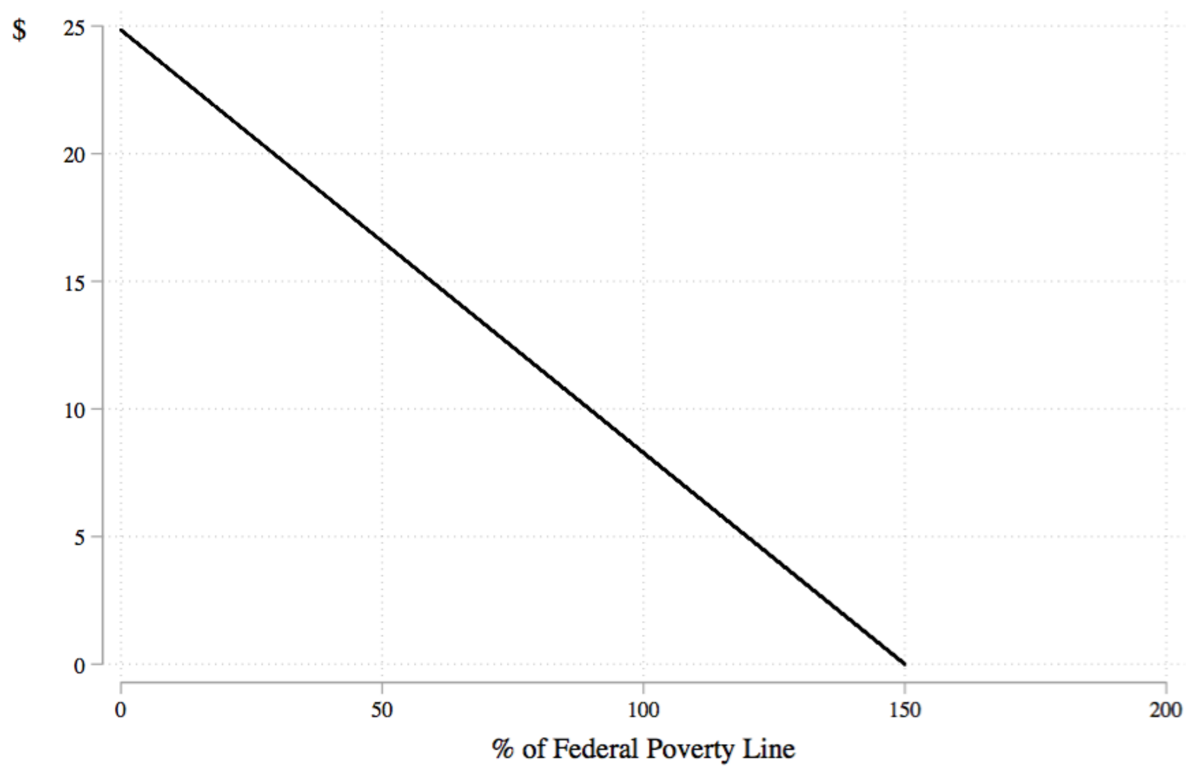
Note: Lines shown are a LOWESS.
Source: Author's calculations using Nielsen Homescan 2016.

Interestingly, fruits and vegetables are a luxury at all income levels but close to the poverty line (100%), the income elasticity trend goes towards 1. This is indicative that at some higher levels of income, we could observe this food group becoming a necessity (income elasticity between 0 and 1). A similar general trend is observed for meats and dairy: they are a luxury at all levels of poverty but go towards 1 the further above the poverty line a household finds itself. A key difference between this food group and fruits and vegetables is that fruits and vegetables are more or less stably a luxury (income elasticity between 1.03 and 1.035) up until close to the poverty line where they experience a downward trend in income elasticity. For meats and dairy, the downward trend is observed for all levels of poverty. Looking back to Figure 2, these two goods exhibited non-linearity in their Engel curves. Therefore, these income elasticities further confirm Banks et al's (1997) findings. An interesting point to consider is that Banks et al's (1997) analysis included more aggregated commodity groups (eg: clothing, transportation) while the analysis presented here shows groups within the "food" category. Thus, we can expect that Engel relationships are complex at various levels of budgeting and non-linearity is to be expect at most of those levels.

5.3. Impact of policy mechanisms

To facilitate the interpretation of the impacts of each policy mechanism, a graphical representation is used. All graphs produced use LOWESS to highlight trends. In order to first establish the impact of the needs-based food assistance program, I plot the gains for each household against the poverty threshold (Figure 4).

Figure 4. Impact of needs-based food assistance.



Note: Line shown is a LOWESS.
Source: Author's calculations using Nielsen Homescan 2016.

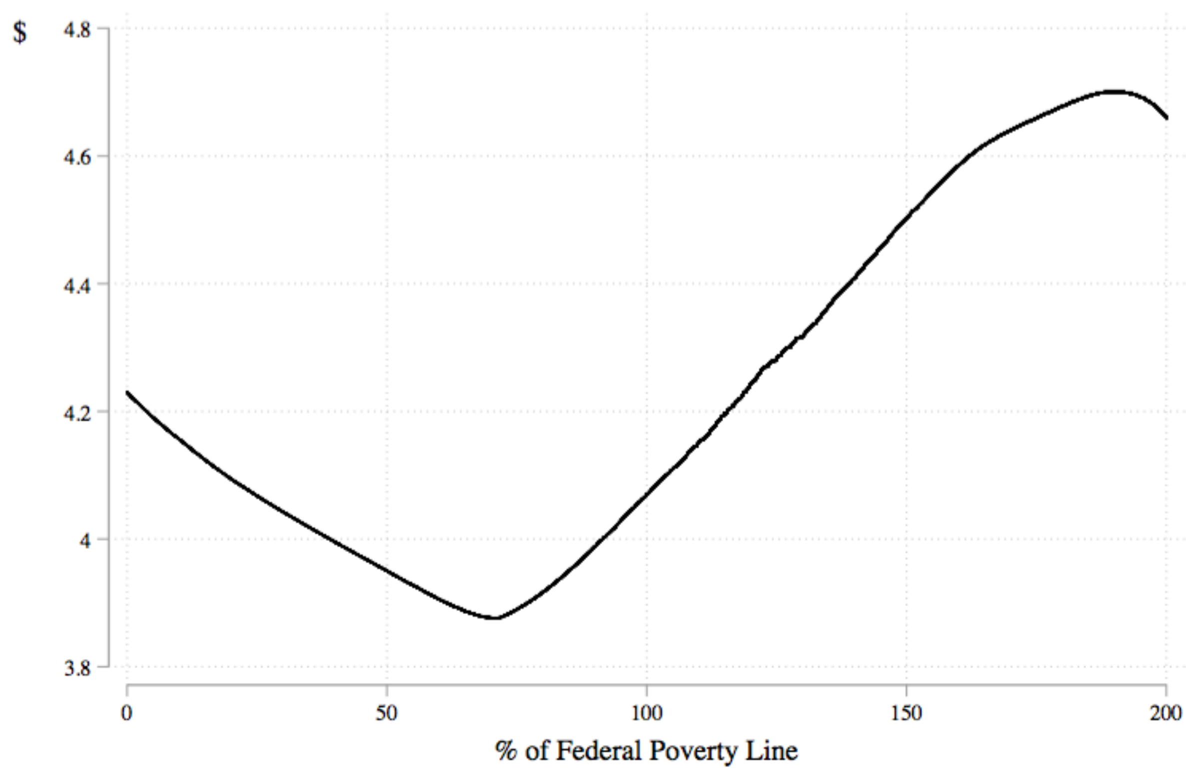
By design, those who are the further below the poverty line (100% on horizontal axis) receive more assistance, while those up to 150% still receive assistance. Those above 150% are not considered for food assistance[2].

The next step is to simulate the impact of a cashback program. In this policy mechanism, households are reimbursed a percentage (~5%) of their spending on fruits and vegetables. Therefore, the gains that a household receives is dependent on how much they spend on fruits and vegetables. It is possible that such a program could stimulate purchases of fruits and

[2] Analysis where those up to 200% of poverty receive food assistance available upon request.

vegetables since households would receive more money the more they buy. It is important to note that this analysis does not account for such incentive effects since it is based on a static environment. Thus, the gains calculated are obtained holding everything constant. The gains of the cashback program can be seen in Figure 5 plotted using a LOWESS.

Figure 5. Impact of cashback program.



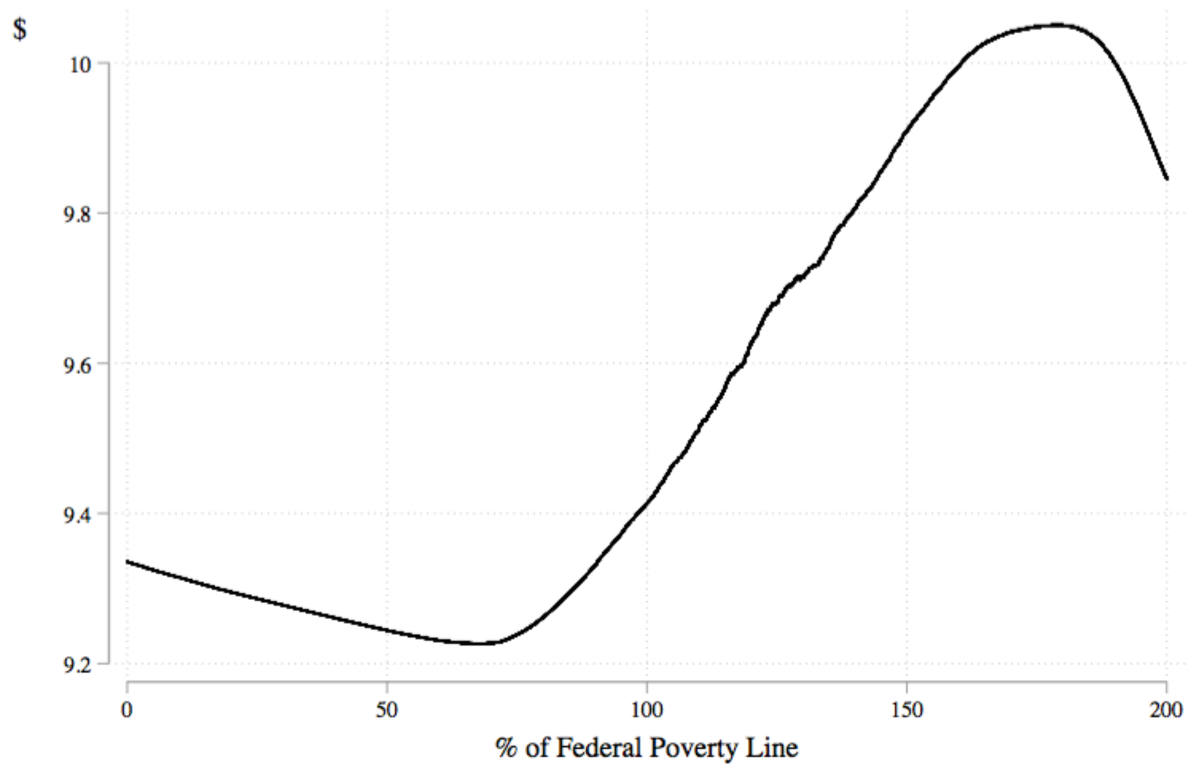
Note: Line shows is a LOWESS.
Source: Author's calculations using Nielsen Homescan 2016.

Interestingly, the gains hover around $3.9-$4.2 for households up to about 75% of poverty. Beyond that point the gains start to rise implying that as households go to lower levels of poverty they gain more from the cashback program. This is likely attributed to those households having the ability to spend more on fruits and vegetables and thus receive more cashback.

The final step is to estimate the impact of a discount on the price of fruits and vegetables. Once again, the analysis calculates the impact of the program holding everything else constant. If such a program were to exist, there might be incentive effects that encourage participating households to increase the purchase of fruits and vegetables. Such effects are unaccounted in the static environment of this analysis.

The impact (CV) of the price discount (~5%) is calculated by using parameters from the QUAI model, as described in Section 3.3. The results are shown in Figure 6 plotted using a LOWESS.

Figure 6. Impact (CV) of discount.

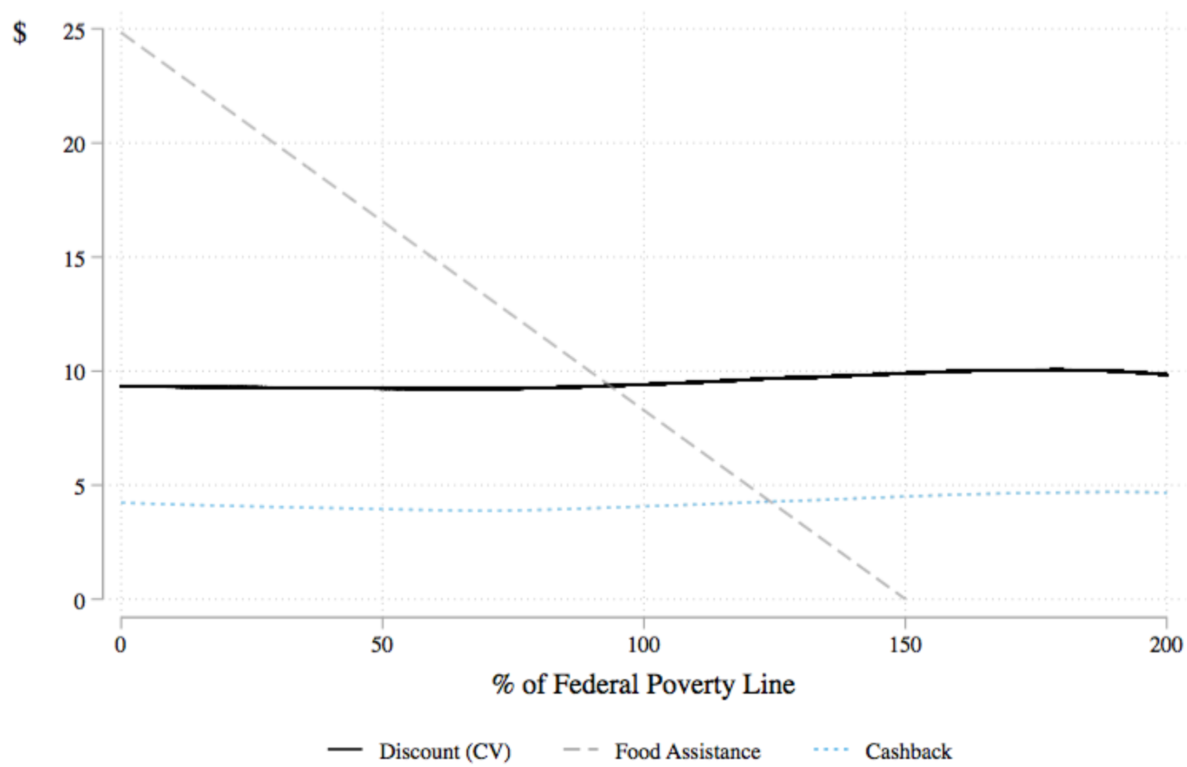


Note: Line shown is a LOWESS.
Source: Author's calculations using Nielsen Homescan 2016.

Similarly to the cashback, the impact of the discount program are somewhat stable (between $9.2 and $9.4) up to about 75% of poverty. From there on, the impact of the program starts to increase indicating that the discount's impacts increase as a household goes to lower levels of poverty. Once again this is likely attributed to these households being able to afford more fruits and vegetables, hence the impact of the discount is larger for them.

In order to allow for a comparison among all three scenarios, Figure 7 shows all of these graphs together on a single plot.

Figure 7. Comparison of all mechanisms.



Note: Line Shown is a LOWESS.
Source: Author's calculations using Nielsen Homescan 2016.

Interestingly, the needs-based food assistance program has a larger positive impact than the proposed mechanisms up until about 100% of poverty. After that point, the discount has a larger positive impact, while the cashback surpasses it at about 125% of poverty. The cashback never surpasses the discount.

6. Discussion and policy implications

An important aspect of food assistance has always been to promote food security and provide much needed leeway to low-income households. In this sense, it is not surprising that needs-based food assistance is currently the most common and most effective food assistance mechanism in the U.S. Looking at the results presented in the previous section, it is clear that this food assistance mechanism is quite helpful for households that are at high levels of poverty. The impact on these households for either of the other two proposed mechanisms are quite small compared to that of needs-based food assistance. This result is quite intuitive: since this households are at high levels of poverty, they are likely not able to purchase enough fruits and vegetables to be able to reap the benefits of either programs.

Conversely, households above the poverty line and around 100% of poverty gain more from the discount. This is likely because they make too much income to receive an amount of needs-based assistance that would be bigger than what they would gain from the other two programs. It is possible that giving these households more money through needs-based assistance could surpass the gains of the discount and cashback program but such an action would require a bigger budget spent on needs-based assistance.

Interestingly enough, the discount on fruits and vegetables provides more gains than the cashback program. It is important to introduce some nuance when comparing these two, however. As discussed earlier, these two programs might have incentive effects on the purchase of fruits and vegetables. In the static environment of this analysis the discount fares better. However, in the real world the size of the incentive effect of these two programs is what would determine which one fares better. If the incentive to purchase fruits and vegetables is higher for the cashback program, it is likely that the gains from it could surpass the gains from the discount. With the current information presented however, the discount fares better.

What does this all mean for public policy? To answer this question, one must consider what policy goals to reach. It is clear that for those who are at or below the poverty line, needs-based food assistance appears to be best. Removing this type of assistance in favor of the other two programs would hurt these households, even if the policy goal is to promote healthy food purchases. Though this analysis does not consider the long-term health impacts of promoting the consumption of healthy foods, it considers the general economic welfare of these households. What good would promoting healthy foods do if the households are worse off? Unless the needs-based assistance is provided in addition to the other mechanisms, this course of action would be quite detrimental.

Conversely, for households who are not at such dire levels of poverty, the cashback and discount program might prove to be beneficial. They would gain more from those programs than from needs-based assistance. Thus, if needs-based assistance is to be slashed in favor of one of the two mechanisms evaluated here, it would only be viable for those who are at various points above the poverty line as they stand to gain from these programs.

Therefore, a potent policy recommendation would be to ensure needs-based assistance for households living at high levels of poverty. If the goal is to promote the purchase of healthy foods, only households beyond the poverty line stand to gain. These programs should in no way replace needs-based food assistance, but instead be used in conjuncture with it to promote healthy food purchases for those who are not living in dire poverty.

7. Conclusion

This paper has been an attempt to compare the impact of two policy mechanisms aimed at promoting the purchase of fruits and vegetables to that of a needs-based food assistance program. The two policy mechanisms in questions are a cashback program where participants are given money back based on their purchase of fruits and vegetables, and the other a discount on the price of fruits and vegetables.

A fictitious budget amount of $1 billion was used to determine the magnitude of these programs. This resulted in a cashback return of about 5% of fruits and vegetable purchases, and a discount of about 5% on fruits and vegetables. The impact for the needs-based food assistance and the cashback program is simply how much money each household would receive from the program based on their current situation. For needs-based assistance, the amount they receive depends on their level of poverty, while for the cashback program it depends on their fruits and vegetables expenditure. To obtain the impact of the discount, I estimate a QUAI model and used the model parameters to calculate CV.

Comparing the impact of all three of these policies reveals that needs-based assistance provides more gains for households at higher levels of poverty, while the cashback and discount surpassed needs-based assistance around the poverty line. Therefore, public policy should prioritize providing needs-based food assistance for households at higher levels of poverty. If encouraging healthy food purchases is a policy goal, households at high levels of poverty could be considered only if they also receive needs-based assistance. Households that are not at high levels of poverty could potentially gain from the discount or cashback program even if they do not receive needs-based assistance.

References

Atkinson, A. B., Gomulka, J., & Stern, N. H. (1990). Spending on alcohol: evidence from the Family Expenditure Survey 1970-1983. The Economic Journal, 100(402), 808-827.

Baltagi, B. H., & Moscone, F. (2010). Health care expenditure and income in the OECD reconsidered: Evidence from panel data. Economic Modelling, 27(4), 804-811.

Banks, J., Blundell, R., & Lewbel, A. (1996). Tax reform and welfare measurement: do we need demand system estimation?. The Economic Journal, 1227-1241.

Banks, J., Blundell, R., & Lewbel, A. (1997). Quadratic Engel curves and consumer demand. The review of economics and statistics, 79(4), 527-539.

Bartlett, S., & Abt Associates. (2014). Evaluation of the Healthy Incentives Pilot (HIP), Final Report. United States Department of Agriculture, Food and Nutrition Service, Office of Policy Support.

Beatty, T. K., Lin, B. H., & Smith, T. A. (2014). Is diet quality improving? Distributional changes in the United States, 1989–2008. American Journal of Agricultural Economics, 96(3), 769-789.

Blundell, R., & Robin, J. M. (1999). Estimation in large and disaggregated demand systems: An estimator for conditionally linear systems. Journal of Applied Econometrics, 209-232.

Blundell, R. W., Browning, M., & Crawford, I. A. (2003). Nonparametric Engel curves and revealed preference. Econometrica, 71(1), 205-240.

Blundell, R., Chen, X., & Kristensen, D. (2007). Semi-Nonparametric IV Estimation of Shape-Invariant Engel Curves. Econometrica, 75(6), 1613-1669.

Browning, M., & Chiappori, P. A. (1998). Efficient intra-household allocations: A general characterization and empirical tests. Econometrica, 1241-1278.

Browning, M., Chiappori, P. A., & Lewbel, A. (2013). Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. Review of Economic Studies, 80(4), 1267-1303.

Burzig, J., & Herrmann, R. (2012). Food expenditure patterns of the generation 50+: an Engel-curve analysis for Germany. British Food Journal, 114(10), 1380-1393.

Dachner, N., Ricciuto, L., Kirkpatrick, S. I., & Tarasuk, V. (2010). Food Purchasing and Food Insecurity: Among Low-income Families in Toronto. Canadian Journal of Dietetic Practice and Research, 71(3), e50-e56.

Davidson, J. E., Hendry, D. F., Srba, F., & Yeo, S. (1978). Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom. The Economic Journal, 661-692.

Deaton, A., & Muellbauer, J. (1980). An almost ideal demand system. The American economic review, 70(3), 312-326.

Dietary Guidelines Advisory Committee. (2015). Scientific report of the 2015 dietary guidelines advisory committee. Washington (DC): USDA and US Department of Health and Human Services.

Dong, D., & Lin, B. H. (2009). Fruit and vegetable consumption by low-income Americans. US Department of Agriculture, Washington, DC.

Drewnowski, A., & Eichelsdoerfer, P. (2010). Can low-income Americans afford a healthy diet?. Nutrition today, 44(6), 246.

Drewnowski, A., & Specter, S. E. (2004). Poverty and obesity: the role of energy density and energy costs. The American journal of clinical nutrition, 79(1), 6-16.

Einav, L., Leibtag, E., & Nevo, A. (2008). On the accuracy of Nielsen Homescan data. USDA, Economic Research Service.

Einav, L., Leibtag, E., & Nevo, A. (2010). Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?. QME, 8(2), 207-239.

Engel, E. (1895). Die Lebenskosten belgischer Arbeiter-Familien früher und jetzt. C. Heinrich.

Fan, L., Baylis, K., Gundersen, C., & Ver Ploeg, M. (2015). Does a Nutritious Diet Cost More in Food Deserts?. In 2015 AAEA & WAEA Joint Annual Meeting, July 26-28, San Francisco, California (No. 205424). Agricultural and Applied Economics Association & Western Agricultural Economics Association.

Jones, E., Akbay, C., Roe, B., & Chern, W. S. (2003). Analyses of consumers' dietary behavior: An application of the AIDS model to supermarket scanner data. Agribusiness, 19(2), 203-221.

Gale, H. F., & Huang, K. (2007). Demand for food quantity and quality in China.

Gregory, C. A., Ver Ploeg, M., Andrews, M., & Coleman-Jensen, A. (2012). SNAP and Diet Quality: An Instrumental Variables Approach. In 2012 Annual Meeting, August 12-14, 2012, Seattle, Washington (No. 124757). Agricultural and Applied Economics Association.

Inglis, V., Ball, K., & Crawford, D. (2009). Does modifying the household food budget predict changes in the healthfulness of purchasing choices among low-and high-income women?. Appetite, 52(2), 273-279.

Leser, C. E. V. (1963). Forms of Engel functions. Econometrica: Journal of the Econometric Society, 694-703.

Lin, B. (2005). Diet quality usually varies by income status. Amber Waves, 3(4), 4-5.

Magana-Lemus, D., Isdorj, A., & Rosson, C. (2013). Welfare impacts of increasing food prices in Mexico: an application of unrestricted Engel curves and LA/EASI demand system. In SAEA Annual Meeting in Its series 2013 Annual Meeting with (No. 143057).

Mokdad, A. H., Marks, J. S., Stroup, D. F., & Gerberding, J. L. (2004). Actual causes of death in the United States, 2000. Jama, 291(10), 1238-1245.

Semega, J. L., Fontenot, K. R., & Kollar, M. A. (2017). Income and Poverty in the United States: 2016. Washington, DC: United States Census Bureau.

Turrell, G., & Kavanagh, A. M. (2006). Socio-economic pathways to diet: modelling the association between socio-economic position and food purchasing behaviour. Public health nutrition, 9(3), 375-383.

United States Department of Agriculture. (2017, December 5). Food Access Research Atlas Documentation. Retrieved July 25, 2018, from https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation/#definitions

United States Department of Agriculture. (2017). Definitions of Food Security. Retrieved September 28, 2017, from https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/definitions-of-food-security/

Wang, D. D., Leung, C. W., Li, Y., Ding, E. L., Chiuve, S. E., Hu, F. B., & Willett, W. C. (2014). Trends in dietary quality among adults in the United States, 1999 through 2010. JAMA internal medicine, 174(10), 1587-1595.

Working, H. (1943). Statistical laws of family expenditure. Journal of the American Statistical Association, 38(221), 43-56.