Low Income and Access to Healthy Food: The Case of Milk

Jeffrey Young
*Murray State University*, jyoung6@murraystate.edu

James K. Binkley
*Purdue University*

Follow this and additional works at: https://digitalcommons.murraystate.edu/faculty

Part of the Agriculture Commons

**Recommended Citation**

This Conference Paper is brought to you for free and open access by Murray State's Digital Commons. It has been accepted for inclusion in Faculty & Staff Research and Creative Activity by an authorized administrator of Murray State's Digital Commons. For more information, please contact msu.digitalcommons@murraystate.edu.
Low Income and Access to Healthy Food: The Case of Milk


Final: June 11th, 2020

Jeffrey S. Young¹ and James K. Binkley²

¹ Department of Agricultural Science (Agribusiness), Murray State University
² Department of Agricultural Economics, Purdue University

Copyright 2020 by the authors. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright appears on all such copies.
LOW INCOME AND ACCESS TO HEALTHY FOOD: THE CASE OF MILK

Abstract

Even with abundant dietary research and knowledge about healthy eating, there is still a considerable shortfall in dietary quality of low-income households in the US. Chen et al. (2012) proposed that this problem is at least in part due to low-income individuals’ having less expected future earnings and thus, reduced incentive to pursue longevity by means of healthy eating. We further their analysis, offering improvement by controlling for access to low-fat milk. We find that even when allegedly crucial factors such as cost and access are not issues, relative purchasing of low-fat milk is still linked to higher income.

Key words: household income, nutrition, access, food cost, milk purchasing.

JEL codes: D12, D15, I12, I14.

Some of the data in this work is calculated (or Derived) based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.
Introduction

The United States Department of Agriculture (USDA) has spent considerable resources on food research including cost, access, availability, preparation, storage, marketing, and nutrition of different kinds of food. In more recent years, there has been an emphasis on studying these subjects as they pertain to healthy eating. Even with the abundance of dietary research and presumably the concomitant knowledge of healthy eating, there is still a considerable shortfall in the nutritional quality of diets, especially of low-income US households (Drewnowski & Specter 2004; Golan et al. 2008). As a result, lower-income individuals have a greater incidence of dietary-related health issues, particularly obesity and diabetes (Robbins et al. 2001; Guo et al. 2004). It remains unclear exactly why this problem persists.

There is a large body of research attributing the link between income and diet quality to the problem of cost (Mooney 1990; Jetter & Cassady 2006; Maillot et al. 2007; Drewnowski, 2010; Monsivais et al. 2011; Rao et al. 2013). Their general conclusion is that low-income households can only easily afford energy-dense foods, which are said to be low cost. A major flaw in many of the of the studies purporting the higher cost of healthy eating is that they often measure food cost as dollars per calorie, or “food energy cost”, which makes very little economic sense. Paraphrasing Chen et al. (2012), energy density has a negative shadow value in a world of excess calories. Moreover, consumers do not tend to make their grocery purchasing decisions based on the caloric or nutrient content of foods, but rather quantity-based criteria such as edible weight and serving sizes (Rolls et al. 2002; Krukowski et al. 2006). In a USDA study, Carlson & Frazão (2012) show that the method used to measure the cost can determine the outcome of the question of whether or not healthier foods cost more. They demonstrate this by calculating the cost of food in three ways, (i) dollars per calorie, (ii) dollars per edible gram, and (iii) dollars per
serving. Only in the first case was healthier food found to be more expensive. A more recent study by Davis & Carlson (2015) showed that the correlation between energy cost and energy density is spurious and energy density does not reduce cost of food. Put differently, the negative relation between energy cost and energy density is necessarily negative – calories in the denominator on one side of the equation and calories in the numerator on the other side of the equation is a mathematical artifact leading to an inverse association. Later work by Drewnowski (2013; 2015) appeared to retreat from measuring food cost in dollars per calorie.

Aside from cost, there is another suspected impediment to healthy eating, which is the issue of unequal access to healthy foods. The term “food desert” has gained popularity as a description of regions or neighborhoods without sources of fresh produce, and the presence of fast food restaurants is widespread. These areas also tend to have a higher concentration of convenience stores, which tend to feature limited, higher-priced selections of healthy foods (Moore et al. 2006). Studies contend that the lack of access to and availability of fresh produce and low-calorie, low-fat varieties of foods force the poor who live in food deserts to buy only what is available, and, hence, their dietary quality suffers as a result (Zenk et al. 2005; Moore et al. 2006; Hilmers et al. 2012). However, the food desert literature as a whole is somewhat conflicted with itself as to the degree that access and availability of healthy food actually matters (Allcott et al. 2019).

This suggests that other factors are at work. Binkley & Golub (2011) argued in favor of a taste-nutrition tradeoff and provided empirical evidence of this using consumer grocery purchasing data. They found a significantly positive relationship between income and purchases of low fat, reduced sugar, and high fiber groceries, after controlling for demographics. That is, consumers with lower income tend to choose the options with higher fat, more sugar, and less
fiber, even though the less healthy options of the foods examined cost either the same or more than the healthier options. Moreover, the commodities chosen were so simply differentiated (e.g. sugar or calorie content in diet versus regular soft drinks) that lack of knowledge about nutrition was unlikely to be an obstacle to choosing the healthier of two options, either.

Chen et al. (2012) furthered this analysis for two specific commodities using food consumption data rather than grocery purchasing data, which allowed them to assess effects of income directly on caloric intake. They used several years’ worth of data taken from the National Health and Nutrition Examination Survey (NHANES) and the Continuing Survey of Food Intakes by Individuals (CFSII) to compute purchasing shares of more healthy types of milk and soft drinks. These were chosen since healthier types are not more expensive, and the types are easily distinguished by the amount of a single nutrient: saturated fat in the case of milk and added sugar for soft drinks. These nutrients are commonly blamed for their contribution to the obesity epidemic in the US\(^1\), while indisputably recognized for their power of taste enhancement. The purchasing shares of healthier types of groceries were found to be correlated with various household and demographic variables. Most important, they found a positive relation in income, that is, low-income consumers did not purchase healthier versions despite the absence of higher price.

This paper also explores the link between income and nutrition by empirically addressing the hypothesis that healthy food costs more, but with an important distinction from previous

---

\(^1\) The Dietary Guidelines for Americans has from its beginning emphasized moderation of consuming fats, sugars, and sodium. Moreover, the 2010 report of the Surgeon General’s office outlines multiple strategies for preventing obesity that concern these nutrients – including recommendations for individuals to reduce “consumption of energy dense foods that primarily contain added sugars or solid fats” as well as “sodas and juices with added sugars”, and choosing to consume more “low-fat or non-fat dairy products” and “fruits, vegetables, whole grains, and lean proteins” (2010). Finally, the Department of Health & Human Services Healthy People 2020 initiative promotes an objective specifically targeting overweight and obesity, including a section which aims at reducing Americans’ consumption of fat and sugar (HP2020, 2017).
work: Chen et al. (2012) could not control for access since their analysis used data from NHANES, which has no information on buying options available to respondents, and Binkley & Gollub (2011) used aggregate market level data. If consumers with lower income have reduced access to healthier types of food, then this could inflate the correlation between income and healthy food consumption, thus compromising these studies. Addressing this gap by controlling for access in the analysis is not only important, but also offers economic insight into whether or not access is a driving factor of the income-nutrition problem. To do so for the case of milk is the chief contribution of this paper.

Milk is an ideal commodity to study. It is a homogeneous commodity that is widely available with fixed container sizes. It is differentiated primarily by fat content, which makes it very easy to determine the “healthiness” of the various choices. For all choices, there is no associated time cost of preparation or other differences in convenience. Further, milk of different fat content is typically sold for the same price, and when this is not the case, price tends to increase with fat content, which we verify with summary statistics in the data section of the study. Thus, the healthier alternative of the two milk types happens also to be the more affordable.

Another benefit to choosing milk is that there is much research examining the access and availability of low-fat milk, some finding evidence of limitations (e.g. Cheadle et al. 1990; 1991; Glanz et al. 2007) and some finding no issues with access (e.g. Hosler et al. 2006; Liese et al. 2007). Consequently, it seems warranted to conduct further research offering further insight into the subject of access to inform policy. By incorporating control of access into the model, we can answer whether or not the findings of Chen et al. (2012) meaningfully change when access is
fixed. If not, then the increasing skepticism of the role of access in linking income to nutrition is justified.

Data

Following the procedure introduced in Chen et al. (2012), we estimate a variation of their econometric model that regresses household purchasing shares of low-fat milk on a host of covariates presented in the original paper. Where our version of this model chiefly differs from the original is in controlling for access to low-fat milk. Doing so required use of different data, which for our case is the AC Nielsen Consumer Homescan Panel, obtained from the University of Chicago Booth School of Business’ Kilts Marketing Center. These data cover a wide geography recording the grocery purchases of more than 40,000 households in US markets. A market is a sampling unit in the Nielsen data – of which there are 52 primary – consisting of a cluster of counties in and around a major metropolitan area. The samples in the markets are representative of the region, and the data includes a projection factor to make any subsamples representative of the US population. The panel includes prices and purchasing sources of more than three million UPC-coded foods as well as household demographic information.

Low-Fat Milk Purchasing

The outcome variable is the quantity share of low fat milk (skim and 1%) of a household’s total gallons of milk purchased during the year 2010. A key difference between low fat and high fat milk is taste, since the primary distinction between nutrient content is saturated fat. Therefore, buying low fat milk involves sacrificing taste in order to gain health by avoiding excess saturated fat. The data allow us to model this variable as a linear function of household
income and demographic variables. Further, because these data record the purchasing sources of milk, imposing restrictions on which data are used allows for controlling access. What is required is that the restricted data pertain to purchases made under conditions where the household could have bought milk of any fat content. This is done by limiting which stores appear as the source in the set of milk purchases made by households—specifically, store chains that offer all types of milk in a given market area. To accomplish this, only stores for which the average price per gallon for all four fat content levels was not missing are included. This was further restricted to store chains which sold at least 100 gallons of each type of milk (skim, 1%, 2%, and whole) during the year in a given market.

**Household Income**

The variable of interest is annual household income. Nielsen measures annual household income by category, in $10,000 to $25,000 increments. This variable was rescaled to the midpoint of the income bracket, making the interpretation of a unit increase in household income to be a $1000/year change rather than a discrete, categorical change. The models from Binkley & Golub (2011) and Chen et al. (2012) that we follow tested the effect of income on the purchasing share of low fat milk. We estimate a slight variant of this model using data constructed in such a way to control for access.

Income correlates positively with the dependent variable, the summary statistics for which are shown in Table 1.
Table 1 Low Fat Milk’s Average Share of Total Milk Purchases by Income Range

<table>
<thead>
<tr>
<th>Income Bracket</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;$20k</td>
<td>0.31</td>
<td>0.43</td>
</tr>
<tr>
<td>&gt;$20k to &lt;$50k</td>
<td>0.36</td>
<td>0.45</td>
</tr>
<tr>
<td>&gt;$50k to &lt;$100k</td>
<td>0.44</td>
<td>0.46</td>
</tr>
<tr>
<td>&gt;$100k</td>
<td>0.53</td>
<td>0.46</td>
</tr>
</tbody>
</table>

The average share of low fat milk of total household milk purchasing is 31% for households whose annual income is under $20,000. For households with annual income over $100,000, 53% of their total milk purchasing is low fat milk.

Milk Prices

In order to control for cost, the average prices for 2010 in dollars per fluid ounce for gallons of skim, whole, 1%, and 2% milk across each store chain for every market was calculated. Doing so allowed the subsetting of the data to include only the store chains in a given market for which the price of each type of milk was not missing, and for which there was at least 100 observations of each type of milk at that chain in the data. Hence, if a consumer purchased milk from that chain and that purchase was recorded in the data, then access to both high fat and low fat milk was not an issue. With the share of low-fat milk as the outcome variable, the coefficient on the price of low-fat milk should be negative and that on high-fat milk should be positive (since these calculations are for shelf prices, not prices paid by consumers including coupons, sales, etc.). We also restrict the data to be purchases of gallons only, which is the most commonly purchased size of milk. In Table 2, it is clear that the average price per gallon of milk increases by fat content.
Table 2 Milk Prices by Fat Content

<table>
<thead>
<tr>
<th>Fat Content</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>0% (Skim)</td>
<td>$0.42</td>
<td>$5.01</td>
<td>$2.08</td>
<td>$0.41</td>
<td>$2.04</td>
</tr>
<tr>
<td>1%</td>
<td>$0.75</td>
<td>$8.04</td>
<td>$2.14</td>
<td>$0.51</td>
<td>$2.11</td>
</tr>
<tr>
<td>2%</td>
<td>$0.87</td>
<td>$6.04</td>
<td>$2.14</td>
<td>$0.42</td>
<td>$2.11</td>
</tr>
<tr>
<td>3.5% (Whole)</td>
<td>$0.82</td>
<td>$4.92</td>
<td>$2.21</td>
<td>$0.44</td>
<td>$2.19</td>
</tr>
</tbody>
</table>

Demographics

Also included are demographic variables that may be correlated with a household’s income as well as its share of low-fat milk purchasing. Race indicators are in the model to help capture differences in milk purchasing by cultural norms and habits, and also because of their correlation with household income. Age and presence of children indicators for infants, youth, and teens are included because consumption of high fat milk tends to be higher for young children. For similar reasons, household size is also included. Educational attainment indicators for highschool and college graduate are included because of the implied increase in knowledge about health and nutrition. Marital status and age bracket (young adult, middle-aged, senior) dummy variables of the head of household are also included, since these tend to correlate with dietary quality (Ervin 2011). Indicator variables for previous or current WIC participation are in the model as well. Aside from their obvious correlation with income, and in addition to milk being in the food package subsidized by WIC, one reason to include the variables is that WIC participants are required to attend nutrition education courses in order to continue receiving benefits. This could skew the household’s purchasing towards low fat milk, at least in the short run. Total ounces of breakfast cereal purchased in 2010 is also included because the taste of milk is less likely to be important when combined with cereal than it is when drinking milk by itself.
Finally, market fixed effects are in the model to help capture geographic differences in milk purchasing.

**Model**

To increase sample validity, we exclude any households who bought less than 10 gallons in 2010, making the final data a sample of 27,437 households in 52 market areas across the US. The model was estimated using the sampling weights provided by Nielsen for each household.

For household $i$, the model takes the form

$$ Y_i = \beta_0 + \beta_1 \ln c_i + X_i \delta + \epsilon_i $$

$Y_i$ is household $i$’s share of low-fat milk, $\ln c_i$ is their annual household income, and $X_i$ is a vector of the other variables. Since, on average, low fat milk is the cheaper option, one could argue that this estimate should be negative: low fat milk is at worst more affordable, and thus should be more attractive to low income consumers.

**Results**

In Table 3 are the results of estimating the model with OLS
**Table 1** Estimation Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.3191***</td>
<td>12.88</td>
</tr>
<tr>
<td>Household Income</td>
<td>0.0011***</td>
<td>16.28</td>
</tr>
<tr>
<td>College Graduate</td>
<td>0.1596***</td>
<td>21.79</td>
</tr>
<tr>
<td>Highschool Graduate, No College</td>
<td>0.0379***</td>
<td>6.45</td>
</tr>
<tr>
<td>Married</td>
<td>0.0417***</td>
<td>6.69</td>
</tr>
<tr>
<td>Young Adult</td>
<td>-0.0192***</td>
<td>-2.85</td>
</tr>
<tr>
<td>Middle-Aged</td>
<td>-0.0128*</td>
<td>-1.84</td>
</tr>
<tr>
<td>Child Age &lt; 2 present</td>
<td>-0.0212***</td>
<td>-2.55</td>
</tr>
<tr>
<td>Child Age 3 to 12 present</td>
<td>0.0421***</td>
<td>6.91</td>
</tr>
<tr>
<td>Child Age 13 to 17 present</td>
<td>0.0223***</td>
<td>2.98</td>
</tr>
<tr>
<td>Black</td>
<td>-0.1615***</td>
<td>-15.63</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.0599***</td>
<td>-7.28</td>
</tr>
<tr>
<td>East US</td>
<td>0.0160</td>
<td>0.93</td>
</tr>
<tr>
<td>South US</td>
<td>-0.1375***</td>
<td>-11.92</td>
</tr>
<tr>
<td>West US</td>
<td>-0.1196***</td>
<td>-8.21</td>
</tr>
<tr>
<td>Price of Low-Fat Milk</td>
<td>-0.0673***</td>
<td>-4.17</td>
</tr>
<tr>
<td>Price of High-Fat Milk</td>
<td>0.0977***</td>
<td>5.85</td>
</tr>
<tr>
<td>Ounces of Cereal Purchased</td>
<td>0.0001***</td>
<td>13.61</td>
</tr>
<tr>
<td>Previously Enrolled in WIC</td>
<td>-0.0455***</td>
<td>-5.95</td>
</tr>
<tr>
<td>Currently Enrolled in WIC</td>
<td>0.0067</td>
<td>0.38</td>
</tr>
</tbody>
</table>

*Note: The R² for the model was about 0.12. Asterisks *, **, and *** represent statistical significance at the p = 0.10, 0.05, and 0.01 levels, respectively.*

From this estimation, it is clear that even when controlling for access, the effect of income is positive and very significant. A $1,000 per year increase in household income coincides with a 0.11% increase in the share of low-fat milk purchased by a household. In Table 1, a discrete jump in household income from <$20,000 to between $20,000 and $50,000 coincided with a 5%
increase in the share, which is comparable the change in Table 3. The signs on the prices correctly reflect the law of demand: when low fat milk becomes more expensive, holding the price of high fat milk constant, consumers buy relatively less low fat milk. The opposite is true for the price of high fat milk. Households with middle aged and young heads purchase less low fat milk than the omitted group, which is senior-headed households, which is consistent with previous findings that dietary behavior is generally better among the elderly (Hann et al. 2001; Ervin 2011). The estimate on cereal indicates that for each additional ounce of breakfast cereal purchased in 2010 by a household, the share of low fat milk purchasing increases by 0.0001. Thus, for each additional 15-ounce box of cereal purchased per year by a household, the milk-purchasing share of low fat milk increases by 0.15\%, which is consistent with the assertion that milk taste is less important when milk is consumed with cereal than when consumed by itself. The effects of the education variables are consistent with prior expectations. The negative slope on the race indicator for a black household may reflect lower expected future prospects and reduced longevity. Therefore, like in the main argument with regards to the effect of income by itself, less value is placed on health as a means to longevity. Current WIC participation coincides with non-decreased purchasing of low fat milk, while previous enrollment is negatively correlated with low fat milk purchasing. One possible explanation for this is that, while enrolled in WIC, some households did shift their milk purchasing towards low fat, but after enrollment ceased and benefits discontinued, once-enrolled households reverted back to purchasing milk with higher fat content.

We offer five observations with regards to this analysis. First, we estimated with only married households and only single households. The results did not differ, particularly with regards to the role of income. Second, the data is extensive enough to run the model using only
data from one chain operating in many markets. There are several such chains. This restriction should control access even more. When this was done, the results differ little from those in Table 3. In particular, the coefficient on income is always positive and highly significant. Third, the model also was estimated using a censored regression. This is because the dependent variable is left-censored at 0 (no low fat milk purchasing) and right-censored at 1 (purchasing is exclusively low fat). The marginal effect of income at the mean was slightly larger than that estimated with OLS. Fourth, the dependent variable was transformed into a categorical variable where 0 indicates no low fat milk purchasing, 1 indicates a share between 0% and 100%, and 2 indicates exclusive purchasing of low fat milk. Then, the model also was estimated using an ordered logit regression. The results are consistent with everything previously presented: a change in income produces a significant increase in the odds ratio of increasing categories of low fat milk purchasing. The fifth and final comment is that the results of this model hold even when the restriction criterion for stores selling all types of milk changes. In the present analysis, each store’s prices for all four types had to be non-missing. Another way to do this is to choose which type had the fewest sales of the four and restrict the number of sales to be greater than some arbitrary threshold. The results for the estimate on income are robust to this alternative as well.

It is clear from the results that in the case of one commodity for which there are no meaningful differences in cost or access but there are differences in taste and nutritional quality, low-income consumers tend to purchase more of the tastier, less healthy version. This outcome is robust to estimation procedure as well as any subsampling or restrictions on the data.
Conclusions

The study herein tests whether access is the cause of the positive association between income and relative healthiness of grocery purchasing. If access is a chief factor in lower-income households purchasing less healthy types of foods (the usual result) then the effect of income in our model should disappear after controlling for access. It does not disappear: those with lower income still choose the less healthy alternative when facing equal access to all types of that food, evidence which implies that the income-nutrition link is not determined by access after all. Because the evidence remains against cost being the driver, and because the explanation of access is weakened, an explanation other than cost or access is required.

One possibility explaining the income-nutrition link’s persistence is that consumers with less income place a lower value on longevity and hence the means to attain it. This idea is not new to economists, but has been explored previously by Grossman (1972), Becker & Murphy (1988), and Binkley (2010). The third study tested this theory by examining the connection between income and smoking, which tends to begin in youth when the length of life is not an issue typically considered (or at least not highly prioritized). However, cost is an issue, so those with lower income are less likely to start. The value of longevity affects the decision to quit smoking, and the author showed that those with lower income are less likely to quit. Exchanging poor dietary choices for smoking, that result is consistent with the outcomes found in Binkley & Gollub (2011), in Chen et al. (2012), and now in the study herein.

This explanation, or the “discounted longevity hypothesis” follows from the point that increasing palatability brings immediate utility and improving nutrition does not (utility of healthy eating is realized in the future, primarily through longer expected life). But the utility of longevity increases with expected income, and is therefore less for individuals with lower
income. As a result, they are less willing to sacrifice the immediate pleasure of tasty eating for the future payoff of healthy eating. This is verified empirically by showing that increased income is linked to a significant reduction of high fat milk purchasing, even though it is, on average, more costly than its equally accessible and healthier alternative.

While this theory has been discussed and tested prior to this work, the contribution of this paper was to control for access in an econometric model of household milk purchasing. The implications are therefore more robustly established than preceding outcomes found by Chen et al. (2012). Because controlling for access in addition to other potentially important variables did not change the effect of income, the theory continues to hold up in light of empirical evidence, at least in the case of milk.
References


Carlson, Andrea, and Elizabeth Frazão. 2012. "Are healthy foods really more expensive? It depends on how you measure the price."


Drewnowski, Adam. 2013. "New metrics of affordable nutrition: which vegetables provide most nutrients for least cost?" *Journal of the Academy of Nutrition and Dietetics* 113(9): 1182-1187.


