HOUSEHOLD FOOD CHOICE IN FOUR FOOD CATEGORIES: HEALTHY OR UNHEALTHY?

by

James Binkley and Alla Golub

Working Paper #10-2

February 2010

Dept. of Agricultural Economics

Purdue University

HOUSEHOLD FOOD CHOICE IN FOUR FOOD CATEGORIES: HEALTHY OR UNHEALTHY?

by
James Binkley
Department of Agricultural Economics, Purdue University
Alla Golub
Department of Global Trade Analysis Project, Purdue University
403 W State Street
West Lafayette, IN 47907
E-mail: jbinkley@purdue.edu; golub@purdue.edu
Working Paper #10-2
February 2010

Abstract

The interest of this study is the tradeoff between nutrition and taste in consumer food choice. We examine choice between more and less healthy versions of particular foods within four important grocery categories: breakfast cereal, milk, bread, and soft drinks. Within a category, products do not greatly differ in terms of cost and convenience, and nutritional differences are easily determined. Consumers are less likely to choose unhealthy foods due to cost or convenience advantages, or through ignorance. This makes the choice between taste and nutrition more apparent. We use annual expenditure data reported by a large sample of households participating in the AC Nielsen Homescan data system. For each of the food categories we develop a measure of 'healthiness' of household expenditures. This is regressed on household demographics and a measure of market prices. We find that households with college-educated heads and higher income households make significantly healthier choices in all the categories. The former was expected. However, although income is generally associated with healthier diets, this is usually attributed to the cost of healthy foods, an explanation inapplicable here. This suggests a deeper understanding of the role of income in diet is needed. As expected, the presence of children leads to a lowering of household nutrition, but primarily for cereal and bread: beverages are little affected. Older households tend to make healthier choices. Finally, we find a reasonably strong role for prices, perhaps reflecting high substitutability among products whose major difference essentially involves a single dimension.

Keywords: Nutrition, Food, Consumer Choice, Health

JEL Codes: D12, I12, I18

Household Food Choice in Four Food Categories:
Healthy or Unhealthy?

In the last few decades considerable evidence has emerged linking food choice to long term health. The USDA has responded by developing nutrition guidelines and otherwise promoting healthy eating. Nevertheless, the average American diet has shown little improvement over the same period. According to USDA data, available calories rose by over 500 between 1970 and 2003, and "even with the mid-1990's push to cut dietary fat, added fats and oils accounted for an extra 216 calories per person per day." (Farah and Buzby, 2005). Recent data from the Centers for Disease Control show that per capita fruit and vegetable consumption falls short of dietary recommendations, having shown no improvement between 1994 and 2005(Blanck et al., 2008). This is the case despite being the food group arguably at the top of most nutrition rankings, and certainly one of the most promoted.

Why do Americans not eat better? There have been many studies attempting to provide insights into this question, especially by relating socioeconomic characteristics-- income, age, race, and education-- to diet quality (Patterson, et al., 1995; Nayga, 1998; Nayga, Tepper, and Rosenzweig,1999; Kennedy, et al., 2000; Kristal, et al., 2001; Wilde, McNamara and Ranney 1999;). At a minimum, these have identified groups with less nutritious eating habits, such as those with low income, and consumers with less education. But reasons for these outcomes can only be inferred. For example, the fact that low income households have less healthy diets is usually attributed to the cost of healthy food (e.g. Drewnowski and Darmon 2005; Golan et al., 2008), while the better diets of educated consumers are often attributed to superior health knowledge (Variyam, Blaylock and Smallwood 1996; Drichoutis, Lazaridis and Nayga 2005).

However, as Blaylock et al. (1999) emphasize, choosing nutritious foods typically involves some sacrifice of taste. Were this not true there would surely be considerably less concern with the American diet. The nutrition-taste tradeoff has grown more pronounced with the growth of food-away from-home and food processing sectors adept at producing foods aimed at the consumer's palate. Increasingly foods feature the addition of fats, sweeteners and salt, all of which enhance palatability but which tend to have negative health effects. Thus,

> In the long run, taste considerations may simply prevail: habits and other forces may be too difficult to overcome. Similarly, the uncertain future benefits of better nutrition—you have to die of something—may outweigh the perceived potential benefits of healthy eating. Put differently, for many people healthy eating is just not worth the effort and sacrifice (Blaylock et al, p 275).

Indeed, consumers tend to rank taste over nutrition (Glanz et al., 1998; Blanck, et al., 2009). Nayga et al. (1999) found that people who ranked taste as important consumed more foods high in fat and sweeteners.

The focus of this paper is an examination of consumer factors affecting the nutrition-taste tradeoff. We attempt to control for the effect of other important influences on food choice by examining household choice between more nutritious and less nutritious versions of particular

food products. Within a narrowly defined product category, differences in cost and convenience are reduced. Then choosing the less nutritious option is more likely to reflect a disinterest in nutrition. Our focus is the extent to which this choice varies by household characteristics, especially household structure—age and presence of children—and socioeconomic status.

We study four product categories: milk, bread, breakfast cereals, and soft drinks. For a focus on nutrition, these foods are excellent products to study. They are each major grocery categories, purchased by nearly all households. Indeed, in the data we use they are the leading products, collectively accounting for more than 15 percent of grocery spending. Most consumers buy from these categories. Furthermore, all are major foods consumed by children as well as adults, which permits examination of how household structure affects choice of product type. This is important, for childhood nutrition and childhood obesity are important issues in food policy (Schwartz et al., 2008).

Finally, and most important for present purposes, each of these categories contains products which most consumers perceive as having significantly different nutritional attributes, with differences relatively easy to determine. Specifically:

a) For cereals, nutritional value tends to vary directly with fiber content and inversely with the percentage of sugar. These characteristics are well publicized. Cereal marketers make frequent use of health claims, whereas sugar-sweetened cereals are often criticized in the popular and academic press, especially regarding their marketing to children.

b) The nutritional value of bread is generally perceived as the extent to which it is made from whole grain flour. Since whole grains tend to be darker in color, dark bread is viewed as healthier than white, and for the most part this is accurate. This makes selecting healthier breads relatively easy for most consumers.

c) In the case of soft drinks, while total avoidance is likely the healthiest choice, calorie free "diet" versions are viewed by most people as nutritionally preferable to regular soft drinks, and for present purposes we regard them as "nutritious." Of course there is no problem in distinguishing between the two varieties.

d) More healthy versus less healthy milk is strictly a matter of the fat content, with skim, or "fat free," being the healthiest. Again, determining the fat content of milk is a matter of glancing at the label, with the fat content often being part of the name of the product.[1] How we measured healthiness for these four categories is described below.

Although "there is no accounting for taste," many people would view the less healthy choices for these categories as the more tasteful choices. This is especially the case with soft drinks and milk, where healthy/unhealthy varieties are distinguished by the extent of a single factor, caloric sweeteners and milk fat, respectively. Most people prefer the taste of whole to skim milk, and regular to diet soft drinks. It is less straightforward for bread and especially for cereals, categories with many varieties and considerable latitude for different ingredients. But while someone choosing bran flakes for breakfast may do it strictly for the taste, it is unlikely anyone chooses chocolate coated sugar bombs for nutrition. In other words, purchasing bran flakes may not be a highly accurate indicator of a demand for nutrition, but purchasing sugar bombs is a good indicator of its absence.

Of course one factor potentially affecting the outcome of the nutrition-taste tradeoff is relative prices. As indicated above, it is often suggested that cost is a significant barrier to a nutritious diet, especially for low income households, because healthy food costs more. However, supporting evidence is somewhat circumstantial and at best associative. For example, it is contended that nutritious foods such as fruits and vegetables are costly, at least in terms of calories,[2] and that low income consumers have less access to such foods and relatively easy access to high calorie fast food. But due to the wide variety of foods available, direct evidence that cost limits healthy choices is lacking. Certainly there are many low priced nutritious foods, as well as high cost foods of dubious nutritional value.

Studying choice within a food category facilitates an examination of income as a limiting factor in healthy food choice. If the cost of healthy food is the primary reason that low income individuals have less healthy diets than do those with higher incomes, then it should follow that, in cases without cost differences, diet choices should be similar across income groups. Cost differences between varieties of a food product are smaller than those for broad food categories, and in the foods we study, healthy choices are not consistently more expensive. Specifically:

a) In the case of milk, most supermarkets follow one of two pricing methods: either all milk types are sold for the same price, or the price increases as the fat content increases.[3] The latter reflects the fact that milkfat is a valuable commodity, being a key ingredient in many manufactured dairy products. Thus, on average, for milk the healthier choices are the lowest in cost.

b) For soft drinks, all products of a given manufacturer in a given store at a given time are almost always priced the same. For this reason, price should have no systematic impact on the choice between regular and diet soft drinks.

c) Because of the large number of cereal manufacturers, each marketing many specific brands, both more healthy and less healthy cereals are each available in low priced and high priced versions. However, the cheapest cereals are those that are relatively unprocessed, such as raisin bran and shredded wheat. Since processing of grains tends to reduce their nutritional value, unprocessed cereals also tend to be healthier. The result is that, *on average*, healthier cereals are cheaper cereals. This is especially true since unprocessed cereals with fewer ingredients are relatively easy to duplicate and thus more prone to competition by cheaper store brands. But because of the diversity of brands available, it does not seem likely that cereal prices are a critical factor affecting the healthiness of cereal choice by income groups.

d) As with cereal, bread is available in a great diversity of kinds and makers, including that produced by in-store bakeries. Since a large proportion of more expensive artisanal and specialized breads tend to use whole grains, the average price of darker breads is likely to exceed that of white breads. But it is not difficult to find both light and dark breads selling for low prices and high prices. This is especially the case since the leading brand in most supermarkets is likely to be the store label, whose standard packaged loaf is available in white or "wheat" (i.e. partially whole wheat) at the same price.[4] So as with cereal, prices are not highly correlated with nutritional values.

Thus, although low income may have a restraining effect on total purchases from these categories, it should not be an important factor determining the type of product within a category. This topic is explored in our analysis.

Data and Methods

The data employed in the study is taken from the 1999 ACNielsen Homescan database, consisting of all retail food purchases (from supermarkets, warehouse stores, convenience stores, drug stores, etc.) and prices paid by 7195 US households during that year. These households represent 52 metropolitan market areas and one area consisting of the rest of the US. The areas are aggregations of counties. Nielsen supplies participating households with home scanning equipment with which they record their purchase data, which is submitted to Nielsen on a regular basis. The data includes the prices and quantities of all food items purchased for home consumption, with complete product information. The database also contains households' demographic characteristics, which are described below. Most of the 7195 households reported purchasing from the four food categories included in this study. Specifically, 7026, 7067, 6998, and 7073 households purchased milk, bread, cereal, and soft drinks, respectively.

Measuring Nutrition

Measuring the healthiness of each household's purchases from the four categories presented varying degrees of difficulty. Two, milk and soft drinks, are straightforward. For milk we used the average fat content for all milk purchased, a number varying between 0% (all skim) and 3.2% (all whole). These were based on product descriptors in the Nielsen data. For soft drinks we used the percent of all purchases comprised of diet varieties. This was especially easy because in the data regular soft drinks and diet soft drinks are separate categories.

Bread was somewhat more difficult. Because of the large number of bread manufacturers, each producing many varieties, there are no descriptive variables neatly dividing bread into well defined nutritional types. As indicated above, we adopted the view that white breads are the less healthy breads. While this is not unerring, the constituent white flour is highly processed, and thus low in whole grains. Furthermore, consumers are likely to judge bread nutrition by this criterion. Using the descriptors available, and a variable describing each individual product,[5] we used our judgement to classify each bread purchased as being either white or dark. We used the proportion of total household purchases comprised of dark bread varieties as the nutrition measure.

Measuring the nutritional quality of cereal purchases was yet more difficult. As in the case of bread, there are many varieties of cereal, but unlike bread, the differences among cereal varieties can be substantial. For example, both products are made from many grains, but bread is dominated by wheat. Cereal is not dominated by any one grain. The most important nutrients affecting cereal nutrition are fiber and sugar, and in assessing a particular cereal these must be weighted in some manner. We used a healthiness index based on a rating developed for 59 major cereals by the Consumers Union *Consumer Reports* magazine (1986). This CU index incorporates sugar and fiber and also protein, fat, and sodium. In order to employ the index, we needed to know these five nutrients for each cereal in the data, and how they were used to compute a value of the index. For the first, we obtained nutritional contents for each brand in the data, using primarily the USDA food nutrition data base. Major cereals appear directly in the USDA data. For others we either used alternative sources (including cereal boxes) or matched

them with USDA cereals. The total number of cereals identified as different by their nutrition content was 451.
Unfortunately, *Consumer Reports* does not state exactly how they computed the indexes.[6] Our solution was to estimate the index prediction function by regressing their original ratings on the protein, sugar, fiber, fat, and sodium content of the cereals they used, using a linear model. The results of the regression are shown in a table in the appendix. What is important is that the R^2 is >.95, with all coefficients highly significant. Positive health effects are associated with protein and fiber and negative effects with fat, sugar and sodium. We used the equation to estimate a CU index for all cereals not covered in the Consumer Reports data, but present in the ACNielsen database, using the estimated coefficients and the nutritional contents of the brands. We computed a weighted average CU index of the cereals purchased by each household, weighting by the ounces of each brand purchased.

Simple statistics for the sample measures appear in table 1. The average for the CU index is approximately that of Kellogg's Corn Flakes, as presented in the original article. In that article the indices ranged from 20 (Captain Crunch) to 71 (Shredded Wheat). For milk, average fat content is very nearly at the midpoint between the minimum and maximum, while the averages for bread and soft drinks are closer to the less healthy end of the nutrition spectrum.

The lower and upper extremes in the table refer to the percent of households who only buy one type of the food indicated. Thus, for example, 8.5 percent only buy whole milk, with slightly fewer buying only skim. For bread and soft drinks, the shares are biased toward the less healthy choices, matching what we see in the means.

Explanatory Variables

The four nutrition measures were used as the dependent variable in regressions on measures of income, education, household structure, and prices. Except for prices, the same independent variables were employed in all equations. Means and variances of all explanatory variables except prices are shown in table 2. Nearly all of these are binaries, making the means sample percents.

Demographics

The demographic variables were constructed from the information in the ACNielsen database. Income is the only non-binary demographic variable and is measured as income per household member. [7] Our expectation about its effect is somewhat open. In a standard demand situation, the importance of income derives from the budget constraint, which limits the amount of a good that can be purchased. Our interest is in the type of a particular food chosen, given that purchases have been made, so the importance of the budget constraint depends upon the extent of differences in relative prices. Higher income permits more expensive types to be chosen, so if anything income should be positively related to the higher priced type. Given the previous discussion about prices for the four food categories, this suggests little if any role for income, at least as it relates to the budget constraint.

However, because we are examining foods which potentially have long term health consequences, income may have an indirect effect, due to the demand for health. This introduces the possibility of a positive relation between income and healthier food choices. This is related to Grossman's work on the demand for health capital (1972). For example, he argues that health investment increases healthy time available for productive activity, which is more valuable for those with higher earning power. Similarly, Becker and Murphy (1988) point out that higher earnings increase the cost of consuming an unhealthy good because any negative effects on productivity cause a greater loss in earnings. In a related literature, Becker and Phillipson (1998) and particularly Davies and Kuhn (1992) study the effects of annuities on longevity, focusing on the fact that receiving an annuity increases the value of living longer, thus encouraging healthier behavior. Having a higher present income increases the likelihood of higher future income, which would be expected to have effects on health demand similar to those of an annuity. .

Reasons such as these might induce a positive relation between income and healthier choices, at least in cases where price is expected to play no role, such as soft drinks.

The most consistent finding across studies dealing with a wide range of health behaviors is a positive association between education and healthier choices (De Walque, 2007; Sander,1995). As indicated earlier, a common explanation for this result is that more highly educated individuals have more information concerning the consequences of unhealthy behavior. For example, Nayga (2000) found that knowledge accounted for the fact that education is negatively related to obesity. However, studies that have corrected for information typically find it only explains part of an education effect, which suggests there must be other reasons (Kenkel 1991). Furthermore, as discussed above, in the cases examined here the differences between more healthy and less healthy products are fairly obvious and well known. It is hard to imagine a consumer not understanding the difference between a regular Coke and a Diet Coke. However, while nearly any consumer may be able to determine the healthier of two products within these categories, more educated consumers are likely to have a better understanding of potential health consequences of a poor diet. In any case, to control for the effects of education on food product choice, we include a set of binary variables measuring the extent of education by household heads. For households with two heads, we used the maximum education level..

Included in the model are six dummy variables indicating the age group of household heads.[8] We expect older people to make more nutritious choices because they have greater concerns about health than do younger individuals, who are more likely to be concerned with taste.

Of special interest is the effect of children on household food choice. Certainly if left to their own devices, children will base their food choices primarily if not exclusively on taste considerations. But children generally do not make the food purchase decisions, so the effect of children depends on the extent to which they are indulged by their parents. Highly health conscious parents may restrict their children to more nutritious foods than they themselves would choose on their own account, in which case the presence of children might increase the nutritional value of the household diet. But it seems likely that such households are the exception, and that children's preferences play a large role on the kinds of products purchased, especially for cereal and bread. The AC Nielsen data provides information on children by

whether there are children under 6, from 6 to 12, and from 13 to 17, and combinations thereof (e.g. whether there are children under 6 and 6 to 12). The actual number is not given. Therefore, we simply used three binary variables indicating whether there are any children falling into the respective age groups.

Three dummy variables describe the nature of household heads: single female, single male, and married couple. The latter is the reference category and is omitted. We do not have strong expectations regarding these variables, although women tend to be more concerned with nutrition than are men. We used two variables to capture time availability: "part time," indicating that at least one household head is working no more than part time, and "not working," indicating that at least one household head works neither full or part time. [9] Note if the latter equals 1, so does the former. Thus, the first measures the effect of having part of the day free, and the second measures the additional effect, if any, of not working at all. Again, we have no strong expectations regarding these variables. However, if anything more time available for household management and shopping would be expected to be conducive to improved household nutrition. Also, time-pressed parents may be more likely to placate their children by giving in to their food preferences.

A final set of indicator variables is included to capture possible racial, ethnic, and geographic differences.

Prices

In addition to demographics, the models for milk, bread, and cereal also contained prices for more healthy and less healthy types of each. (Prices were not used in the soft drink equation for reasons stated earlier: all products—regular and diet—by a given manufacturer are priced the same.) The Nielsen data has prices paid for all individual purchases. One approach would be to use these prices as paid by each consumer. However, there are practical and statistical difficulties associated with this. One is the problem of obtaining a price for items never purchased: for a household buying only healthy choices in a category, there is no price information for less unhealthy items. Even if there were, however, using prices actually paid opens the way for problems of endogeneity. For example, consumers with low search costs and access to many outlets will search for the lowest price available, especially for frequent purchases. Then the large quantity is more the cause of the low price, rather than vice-versa.[10]

We thus based our prices on area prices. As stated above, ACNielsen data are taken from 53 markets, 52 of which are a cluster of counties around a major city, composing a well-defined metropolitan market. [11] We assumed that all households in the same market faced the same price for healthy products and unhealthy products.[12] For bread and milk, defining healthy and unhealthy product prices followed our definitions. For example, the price for less healthy bread in a market was computed by summing the ounces of white bread purchased by Nielsen households in that market and the dollars paid, and then dividing the latter by the former. The same was done for milk.

To calculate prices for cereal, the healthiness measure is the CU Index, which is continuous. Thus we needed to define "most healthy" and "least healthy" cereals, a definition which is

necessarily arbitrary. We used the upper and lower quartiles of the CU distribution. For each market we calculated the average price cereals falling into the upper and lower quartiles of the respective index, again by dividing total expenditures by total ounces.

Simple statistics for prices (cents/ounce) are shown in table 3. For cereal and milk, the average price for less healthier products is higher. The difference for cereal appears large. This is perhaps deceptive, however, for there is great diversity within both groups of cereal: one can easily find low-cost and high cost cereals of each type. For bread and milk, the differences are small, particularly for the latter. Indeed, in nearly one half the sample markets (44%), the average for less healthy milk was cheaper. Nevertheless we stand by our view that in the vast majority of supermarkets it is either the converse or all milk is priced the same. These prices are estimates, based on the buying behavior of a sample of households, in some cases not a large sample, so they depend on contingencies like store shopped, whether the particular household uses private labels, and so forth. We must assume that they are reasonably representative.

Methodology

We estimated the four models with OLS. However, there are some methodological issues. First, to avoid any possibility of heteroskedasticity-induced bias, t statistics are based on White's heteroskedasticity consistent variance estimator.[13] These have the advantage of being consistent and generally unbiased whether or not there is heteroskedasticity, while avoiding the problem of trying to model it if it is found. This approach is becoming the recommended procedure with cross-section data, especially when samples are large (Wooldridge 2006).

Second, the ACNielsen data is not a random sample. Although there are survey weights associated with the data, it is not clear how they are to be used. Perloff and Denbaly (2007) address this in a recent paper discussing problems associated with using scanner data in research. Fortunately, for this study the issue is somewhat moot. We employed a test for weighting suggested by DuMouchel and Duncan (1983), which failed to reject the hypothesis that weighting was not needed.

Finally, we performed some data transformations to improve the interpretability of the results and their comparability across models. For milk, rather than using the fat content of household milk purchases directly, we subtracted that from 3.2% (the maximum possible). This makes increasing values an indicator of improved nutrition, matching the other three models. Second, we scaled the dependent variables, the prices, and the income variable by subtracting the mean and dividing by the standard deviation. As a result, the coefficients on the indicator variables (i.e. all but prices and income) measure the standard deviation change if the characteristic is present relative to the reference category. For income and prices, the coefficient measures the standard deviation change in the dependent variable due to a one standard deviation increase in income. This transformation increases the comparability across models and variables.

Results

The results of the regressions are shown in table 3. Coefficients with t-statistics exceeding 2 in absolute value are in bold type. For the sample sizes involved, this indicates a significance of .05 or less. The R-squared's, though somewhat low, are typical for data of this type.

We begin with income. In all four cases, increasing income is associated with choices having better nutrition. All income coefficients are significant at .05 or better. The effect is strongest—both statistically and practically-- for milk and soft drinks, for which a one standard deviation income increase (which is 18.45 thousand dollars) is associated with a .15 standard deviation increase in the nutrition measure. These are the two products for which the nutrition-taste tradeoff is strongest: most people prefer the taste of the less healthy product. In addition, the less healthy product is unlikely to be cheaper. This is evidence that viewing the low income-poor diet nexus as a matter of affordability is at best incomplete. It may have more to do with the relation between income and the demand for health, as discussed above.

As expected, education is found to have an important influence on the type of food purchased. However, the effect appears to be confined to college education, which in all four cases is significantly and positively associated with healthier choices. The effect for milk is especially strong. Relative to a high school dropout (the reference group), the coefficient on college implies about 5 less calories for every 8 ounces of milk consumed.[14] The only other case of significance for education is the high school variable in the bread equation. We regard this as a type I error, however, for the sign is clearly "wrong." The importance of college most likely reflects a better understanding of the role of nutrition in health rather than greater nutritional knowledge. It hardly seems reasonable that college graduates have an advantage over, for example, high school graduates in determining the fat content of milk.

The 'single male' and 'single female' variables compare single head households to households with married couples. The coefficient for single females is all positive, but only soft drinks is significant. Single men make significantly poorer beverage choices, and significantly better bread choices. If nutrition is the primary motivation, this is inconsistent. We conclude that nutrition is not the motivation. But if taste is the dominating influence, the outcome need not be inconsistent, for many adults prefer darker breads. Evidently men are especially likely to do so. On this interpretation, the results suggest that women are more concerned about nutrition than are men, as has been suggested by previous research (e.g. Block, et al., 1988).

Age is generally associated with better nutrition choices. With the exception of milk, this is what is suggested by the results: increasing age is associated with better nutrition choices. The effect is strongest for cereal, where the difference between the oldest and youngest age group, which is the reference, is nearly a standard deviation. Why milk differs from the other categories cannot be determined from this data. The pattern of coefficients suggests the result is primarily due to a greater acceptance of low fat milk by the youngest households.

The variables measuring time availability of heads of household are estimated to be of no importance in any model. We do not find this surprising. It is consistent with our argument that the nutrition information needed to make healthy choices for these food types is readily available

and easily gotten. Furthermore, convenience is not an issue when choosing varieties within these categories.

As expected, the results indicate that the presence of children leads to food purchases of reduced nutrition. The effect is strongest for cereal, a category featuring products designed to appeal to children and typically supported by considerable marketing to them. It is also strong for bread. Although the presence of very young children is estimated to lead to higher fat milk, this can be attributed to the recommendation that very young children should be given whole milk to provide extra calories for growth. Overall, children have much smaller impacts on milk and soft drink purchases than on the grain-based foods. This suggests that home consumption of beverage calories may not be a major factor in childhood obesity.
In an expanded version of the model, we investigated whether the coefficients on the children variables varied by household type, specifically, when two parents are present and if head(s) possessed a college degree. This was done by adding interaction terms between the three children variables and the college degree dummy and an indicator for married couples.[15] If anything, we expect such households to do a better job of ensuring that their children maintain a nutritious diet. However, we found no consistent effects. None of the married couple interactions in any of the models was significant at .05. In the case of college, the results suggest that more educated parents actually do a worse job of monitoring nutrition. Half the coefficients were negative. For two of these, teens in the bread and soft drink equations, the t statistics were as large as 2. Although we doubt this reflects a tendency to greater laxness by educated parents, it certainly does not support a belief that they would impose higher nutrition standards.

Regarding racial/ethnic effects, there are some significant differences with whites, the reference group. Hispanics are estimated to be more likely to choose healthier options when purchasing milk and soft drinks. Patterson et al. (1995) found that Hispanics had diets lower in fat than did other groups. African Americans show significant differences in all four cases. For the grain-based foods, the choices are better nutritionally; for beverages they are very much worse. The strength of the latter suggests that the former are more a reflection of taste preferences than concern over nutrition. This interpretation is consistent with the fact that African American diets have generally been found to be inferior nutritionally to those of other groups (Gates and Macdonald, 1997; Patterson et al., 1995).

The final variables are prices. Despite the use of somewhat approximate measures, in all cases the signs is correct, and in five of the six (the exception being the price of healthier cereal) the coefficients are highly significant. These results suggest there may be many consumers on the margin between taste and nutrition. Taxes/subsidies may be an effective means of inducing such consumers to select healthier versions of food products when shopping. For example, a 5 percent increase from the average less healthy milk price is estimated to lead to a .35 of a percent reduction in fat content of milk purchases.[16] Price manipulations within food categories may be particularly effective in affecting choice since various versions are usually shelved together, which facilitates price comparisons, and the different types are likely to be more substitutable than foods from broadly different categories.

Conclusions

In this study we examined household product choice within four important grocery categories: milk, bread, breakfast cereals, and soft drinks. These are interesting because nearly all households purchase from each of them, and each contains varieties with salient differences in nutritional quality. Furthermore, price differences across varieties are small or nonexistent, so that higher cost is unlikely to be a barrier limiting access to healthier products. We used 1999 data from the AC Nielsen Homescan data base, which contains detailed household purchase information of individual grocery products as well as characteristics of the sample households.
For each category we developed a measure of the extent to which a household's purchases showed a tendency to choose healthier versions of the products contained therein. We regressed these on characteristics of the household heads, variables measuring household structure, and market prices.
In several cases the results showed expected patterns. Children were found to shift household purchases to less healthier types of the food categories examined, with stronger effects for bread and cereal than for the two beverages. The nutritional quality of bread and cereal purchases increased sharply with age of household heads. Certainly that is no surprise, but a similar outcome occurred for soft drinks, with older households much more likely to substitute diet for regular soft drinks. However, this was not the case with milk. There were significant ethnic and racial differences, but no conclusive evidence that purchases by any group are consistently higher or lower in nutritional value.

The only household characteristics with effects that were statistically significant and consistent in direction for all four food types were college education and income, both with strong positive effects. That college education is associated with better nutrition is no surprise, for nearly all studies of health behavior have found a similar result. This generally accepted reason for this is greater knowledge of nutritional attributes and a better understanding of the role of nutrition in future health. Studies have also found that increasing income is associated with better nutrition. The reason for this is not as clear, but it is typically attributed to the cost of a healthier diet. However, in this study such an explanation is not plausible. Price differences of more healthy and less healthy products within the four categories are small. Indeed, income had its largest effect for milk and soft drinks, although lower fat milk tends to be cheaper than whole milk and regular and calorie free soft drinks are priced the same. Hence an important conclusion of this study is that an explanation of why lower income consumers have lower nutrition diets may need to go beyond food cost and food access. We believe it may be due to a reluctance to forego the pleasure of favorite foods in return for increased probability of future healthy time. Healthy time is more valuable when income is available to increase the enjoyment of that time. This is a topic for future research.

Finally, three of the four models (soft drinks being the exception) included a representative price for more healthy and for less healthy product varieties. Prices were based on averages over the households in each Nielsen market. All estimated effects had the correct sign, and five of the six were highly significant. The magnitude of the estimates suggests that price changes within these categories can potentially cause demonstrable product switching. This should not be

surprising, for products within a category are necessarily similar and we would expect them to be relatively substitutable.

References

Becker, G. S., and K. M. Murphy. 1988. "A Theory of Rational Addiction." *Journal of Political Economy*, 96 :(August) 675-700.

Becker, G. S., and T. Philippson. 1998. "Old Age, Longevity, and Mortality Contingent Claims." *Journal of Political Economy* 106 551-73.

Binkley, J.K. 2006. "The Effect of Demographic, Economic, and Nutrition Factors on the Frequency of Food Away from Home," *Journal of Consumer Affairs* 40(2): 372–391.

Binkley, J. K, J. Eales, and M. Jekanowski. 2000. "The relation between dietary change and rising US obesity." *International Journal of Obesity* 24: 1032-39.

Blanck, H., C Gillespie, J Kimmons, J. Seymour, and M Serdula (2008). "Trends in Fruit and Vegetable Consumption Among U.S. Men and Women." Preventing Chronic Disease, April 5(2). Published online, available at http://www.cdc.gov/pcd/issues/2008/apr/07_0049.htm, accessed October 6, 2009.

Blanck, H. A. Yaroch, A. Atienza, S. Yi, J Zhang, and L. Masse (2009). "Factors Influencing Lunchtime Food Choices Among Working Americans." *Health Education Behavior*, 36(2), April, 289-301.

Blaylock, J., D. Smallwood, K. Kassel, J. Variyam, and L. Aldrich. 1999. "Economics, Food Choices, and Nutrition" *Food Policy*, 24 :269-286.

Block, G., W. Rosenberger, and B. Patterson. 1988. "Calories, Fat and Cholesterol: Intake Patterns in the US Population by Race, Sex, and Age." *American Journal of Public Health*, 78:9 (September, 1988). 1150-1155.

Carlson, K. A., and B. W. Gould. 1994. "The Role of Health Knowledge in Determining Dietary Fat Intake." *Review of Agricultural Economics*, 16: 373-86.

Cash, S. B., D.L. Sunding, and D. Zilberman. 2004. "Fat Taxes and Thin Subsidies: Prices, Diet, and Health Outcomes." Selected Paper, American Agricultural Economics Association Annual Meeting, Denver, Colorado, USA.

Consumer Reports (1986), "Ready-to-Eat Cereals", 51(10):October 628-637.

Davies, J., and P. Kuhn. 1992. "Social Security, Longevity, and Moral Hazard," *Journal of Public Economics*, 49(1): 91-106.

De Walque, D. 2007. "Does Education Affect Smoking Behaviors? Evidence Using the Vietnam Draft as an Instrument for College Education" *Journal of Health Economics* 26: 877-895.

Drewnowski, A., and N. Darmon. 2005. "The Economics of Obesity: dietary Energy Density and Energy Cost." *American Journal of Clinical Nutrition* 82(1): 265S-273S.

Drichoutis, A.C., P. Lazaridis, and R. M. Nayga. 2005. "Nutrition knowledge and consumer use of nutritional food labels." *European Review of Agricultural Economics* 32(1): 93-118.

DuMouchel, W. H., and G. J. Duncan. 1983. "Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples." *Journal of The American Statistical Association*, 78(383): 535-543.

Farah, Hodan, and Jean Buzby (2005), "U.S. Food Consumption Up 16 Percent Since 1970." *Amber Waves*, November, p.5

Gates, G., and M. McDonald, 1997. "Comparison of Dietary Risk Factor in African-American and White Women." Journal of the American Dietary Association, 97: (12), 1394-1401.

Glanz, K., M. Basil, E. Maibach, J. Goldberg, and D. Snyder. 1998. "Why Americans Eat What They Do: Taste, Nutrition, Cost, Convenience, and Weight Control Concerns as Influences on Food Consumption." *Journal of American Dietetic Association* 98: 1118-1126.

Golan, E., Stewart, H., Kuchler, F., & Dong, D. (2008). Can Low-Income Americans Afford a Healthy Diet?-*Amber Waves, 6* (5), 26-33.

Grossman, M. (1972). "The Concept of Health Capital and the Demand for Health." *Journal of Political Economy*, 80(2), pp 223 -255.

Hendel, I, and A. Nevo. 2006. "Measuring the Implications of Sales and Consumer Inventory Behavior." Econometrica 74(6): 1637-1673.

Jacobson, M. F, and K. D. Brownell. 2000. "Small Taxes on Soft Drinks and Snack Foods to Promote Health." *American Journal of Public Health* 90: 854-57.

Kenkel, D. 1991. "Health Behavior, Health Knowledge, and Schooling." *Journal of Political Economy* 99(2):287-305

Kennedy, E., Bowman, S.A., Lino, M., Gerrior, S.A. and Bisiotis, P. 1999. Diet Quality of Americans. *America's Eating Habits: Changes and Consequences, Agricultural Information Bulletin No 750: 97-109.*

Kinnucan, H. W., H. Xiao, C.-J. Hsia, and J. D. Jackson. 1997. "Effects of Health Information and Generic Advertising on U.S. Meat Demand." *American Journal of Agricultural Economics* 79: 13-23.

Kristal, A. R., Hedderson, M. M., Patterson, R. E., & Neuhauser, M. L. (2001). Predictors of self-initiated, healthful dietary change. *Journal of the American Dietetic Association, 101,* 762–766.

Kuchler, F., A. Tegene, and J. M. Harris. 2005. "Taxing Snack Foods: Manipulating Diet Quality or Financing Information Programs?" Review *of Agricultural Economics*, 27(1): 4-20.

Moulton, B. R., and W. C. Randolph. 1989. "Alternative Tests of the Error Components Model." *Econometrica* 57(3): 685 – 693.

Nayga, R. 1998. Consumer Characteristics Associated with Low Fat, Low Cholesterol Foods. *International Food and Agribusiness Management Review, 1*: 41-49.

Nayga, R., Tepper B. J., and Rosenzweig, L. 1999. Assessing the importance of Health and Nutrition Related Factors on Food Demand: A Variable Preference Investigation. *Applied Economics*, *31*: 1541-1549.

Nayga, R.M. Jr. 1997. "Impact of Sociodemographic Factors on Perceived Importance of Nutrition in Food Shopping.," *Journal of Consumer Affairs*, 31(1): 1-9.

Patterson, B.H., l Harlan, G. Block, and L. Kahle, 1995. "Food Choices of Whites, Blacks, and Hispanics: Data from the 1987 National Health Interview Survey." *Nutrition and Cancer*,23(2): 105-19.

Perloff, J., and M. Denbaly. 2007. "Data Needs for Consumer and Retail Firm Studies" *American Journal of Agricultural Economics* 89(5): 1282-1287.

Putnam, J., and J. Allshouse. 1999. "Food Consumption, Prices, and Expenditures, 1970-97." Food and Rural Economics Division, Economic Research Service, U.S. Department of Agriculture. Statistical Bulletin No. 965.

Sander, W. 1995 "Schooling and Quitting Smoking." *Review of Economics and Statistics* 77(1): 191-199.

Schwartz, M.B., L.R. Vartanian, C.M. Wharton, and K.D.Brownell. 2008. "Examining the Nutritional Quality of Cereals Marketed to Children." *Journal of the American Dietetic Association*, 108 (6): 702-705.

Stoker, R., and J. Rourke, "Retail Prices for Whole Milk and Reduced Fat (2%) Milk 2001-2005 Summary." Dairy Program, Agricultural Marketing Service, U.S. Dept of Agriculture, June 7, 2006. Available at http://www.ams.usda.gov/AMSv1.0/getfile?dDocName=STELDEV3023056 (accessed 8/7/09

Variyam, J., J. Blaylock, and D. Smallwood (1996). "A Probit Latent Variable Model of Nutrition Information and Dietary Fiber Intake." *American Journal of Agricultural Economics*, 78(August), 628-639.

Wilde, P.E., P.E. McNamara and C. K. Ranney. 1999. "The Effect of Income and Food Programs on Dietary Quality: A Seemingly Unrelated Regression Analysis with Error Components." *American Journal of Agricultural Economics* 81(4): 959-71.

Wooldridge, M.. 2006. *Introductory Econometrics: A Modern Approach*, Third Edition. Thomson Southwestern Publishing: Mason, Ohio.

Table 1. Statistics for Health Measures.

Food Measure	Mean	Standard Deviation	Percent at Lower Extreme	Percent at Upper Extreme
Cereal – Avg. CU Index	44.69	9.04	-	-
Milk - Avg Percent Fat	1.57	1.03	7.6	8.5
Bread – Proportion Dark	.389	.348	16.3	4.0
Soft Drinks – Proportion Sugar Free	.323	.352	3.1	25.8

Table 2. Statistics for Explanatory Variables

Variable	*Mean*	*Std Dev*
INCOME	26.148	18.451
HIGH SCHOOL GRAD	0.146	0.353
SOME COLLEGE	0.603	0.489
COLLEGE GRAD	0.474	0.499
SINGLE FEMALE	0.232	0.422
SINGLE MALE	0.097	0.295
TIME1	0.493	0.500
TIME2	0.375	0.484
HH Heads 26-35	0.072	0.258
HH HEADS 35-39	0.107	0.309
HH Heads 40-44	0.131	0.337
HH heads 45-49	0.152	0.359
HH heads 50-54	0.144	0.351
HH Heads 55-64	0.198	0.399
HH HEADS > 64	0.162	0.369
CHILD < 6	0.104	0.306
CHILD 6-12	0.168	0.374
CHILD 13-17	0.151	0.359
AFRICAN AMERICAN	0.117	0.321
EAST ASIAN	0.015	0.123
HISPANIC	1.933	0.250
SOUTH	0.387	0.487
WEST	0.190	0.393
EAST	0.220	0.414

Table 3. Regression Results for Four Food Categories

	BREAD		MILK		CEREAL		SOFT DRINKS	
	BETA	T	BETA	T	BETA	T	BETA	T
CONSTANT	**-0.279**	-2.21	**-0.269**	-2.17	**-0.654**	-5.94	**-0.432**	-3.79
INCOME	**0.041**	2.53	0.151	9.18	**0.055**	3.43	**0.146**	8.82
HIGH SCHOOL GRAD	**-0.130**	-3.53	-0.034	-0.86	0.025	0.64	0.020	0.52
SOME COLLEGE	-0.007	-0.20	0.056	1.65	-0.010	-0.30	-0.042	-1.23
COLLEGE GRAD	**0.229**	6.49	**0.329**	9.43	**0.108**	3.20	**0.077**	2.26
SINGLE FEMALE	0.046	1.41	0.003	0.08	0.050	1.46	**0.083**	2.47
SINGLE MALE	**0.201**	3.88	**-0.212**	-4.21	-0.025	-0.49	**-0.249**	-5.23
TIME1	0.036	0.90	0.061	1.57	0.020	0.53	-0.029	-0.73
TIME2	0.033	0.80	-0.057	-1.38	0.029	0.73	-0.020	-0.49
HH Heads 25-35	**0.203**	2.36	**-0.155**	-2.01	0.093	1.25	0.053	0.71
HH HEADS 35-39	0.130	1.59	**-0.194**	-2.65	0.131	1.86	**0.170**	2.37
HH Heads 40-44	**0.216**	2.68	**-0.171**	-2.35	**0.267**	3.75	**0.226**	3.18
HH heads 45-49	0.144	1.82	-0.139	-1.94	**0.453**	6.37	**0.259**	3.70
HH heads 50-54	**0.217**	2.72	**-0.167**	-2.31	**0.500**	6.96	**0.264**	3.72
HH Heads 55-64	**0.278**	3.56	-0.076	-1.09	**0.693**	9.88	**0.467**	6.68
HH HEADS > 64	**0.409**	5.00	-0.013	-0.17	**0.890**	11.86	**0.453**	6.12
CHILD < 6	-0.072	-1.55	**-0.292**	-7.05	-0.035	-0.97	-0.014	-0.35
CHILD 6-12	**-0.131**	-3.53	-0.061	-1.75	**-0.325**	-10.51	-0.030	-0.89
CHILD 13-17	**-0.217**	-6.07	0.028	0.79	**-0.301**	-9.13	**-0.088**	-2.60
AFRICAN AMERICAN	**0.353**	8.04	**-0.539**	-13.66	**0.173**	4.11	**-0.489**	-14.70
EAST ASIAN	-0.094	-0.89	-0.004	-0.04	0.055	0.50	**-0.273**	-2.89
HISPANIC	-0.029	-0.58	**0.226**	4.56	0.058	1.33	**0.145**	3.21
SOUTH	0.009	0.23	**-0.227**	-6.20	**0.136**	4.33	**-0.087**	-2.55
WEST	0.026	0.63	**-0.115**	-2.67	0.048	1.12	-0.067	-1.64
EAST	**-0.102**	-2.52	**-0.112**	-2.65	0.041	1.06	-0.071	-1.87
LESS HEALTHY PRICE	**0.057**	2.69	**0.068**	3.41	**0.051**	3.31	-	-
MORE HEALTHY PRICE	**-0.104**	-4.92	**-0.087**	-4.32	-0.013	-0.90	-	-
R^2	.080		.127		.168		.092	

Bold indicates significant at ≤ .05.

Appendix Table 1. Weights for Each Component in the CU Index

Variable	Coefficient	Standard Error	t Value	Pr > |t|
Intercept	63.21	2.60	24.28	<.0001
Protein	0.88	0.18	4.89	<.0001
Total lipid (fat)	-0.93	0.15	-6.29	<.0001
Sugars	-0.49	0.04	-12.69	<.0001
Fiber	0.26	0.05	5.41	<.0001
Sodium	-0.02	0.00	-13.92	<.0001

Number of observations	67
Adj R – square	0.94
F statistics	211.41

Note: Dependent variable is Consumer Report Magazine Index.

Footnotes

[1] As reported in Blaylock et al. the 1989-90 USDA Diet and Health Knowledge Survey asked several questions comparing the healthiness of two foods. The most frequently answered correctly were "Which has more fiber, whole wheat bread or white bread?" (92.9%) and "Which has more fat, skim milk or whole milk?" (95.5%).

[2] Although this would seem to make a vice from a virtue. It also ignores the fact that usually the problem is an excess of calories, making their shadow value zero at best.

[3] See, for example, Stoker and Rourke (2006).

[4] In our data over a third of bread purchases were store brands.

[5] This variable is composed primarily of abbreviated terms such as "whi" (white) and "whe" (wheat) which facilitated our classification.

[6] We contacted Consumers Union but they were unable to supply their methodology.

[7] In the ACNielson database, the household income variable is given by intervals. For example, $12 000 - $14 999, $15 000 – $19 999 and so on. Income per household member is calculated as the mean point of the interval divided by size of household.

[8] The ACNielsen measures were not actual ages but intervals. In the case of two household heads in different groups, we took the average, assigning the higher interval if the average was on an interval boundary.

[9] The ACNielsen measures of employment are non-standard: less than 30 hours/week; 30-34 hours; 35 or more hours; and not working. Given these choices, we defined "part-time" as less than 30, and full-time as greater than 30.

[10] Hendel and Nevo (2006) make a similar point regarding temporary price reductions for storable goods.

[11] The remaining ACNielsen region is "other," consisting of mostly rural counties not close to any major markets. It is the largest region, with 1127 households. Since these are scattered throughout the US, it would make little sense to view them as having a representative price. We thus were forced to remove them from the sample. This left us with 5892 households in our final sample.

[12] While certainly prices within a market vary across stores and perhaps areas of cities, we expect greater variation across cities than within a given city.

[13] In fact, Breusch-Pagan tests show the presence of weak heteroskedasticity.

[14] From table 1, the standard deviation of the milk variable is about 1percent fat content, which amounts to about 15 calories per 8 ounces. The coefficient on college is one third of that. Based

on a 20 gallon annual per capita milk consumption, this is a difference of 1600 calories a year on this single food item.

[15] These results are available on request.

[16] A 5 percent increase is .05 x 2.14=.107, which is .73 standard deviations. From the results, a one standard deviation change in price causes a .484 standard deviation change in fat content, so a .73 standard deviation change causes a .355 standard deviation in fat content. The standard deviation of fat content is 1.034. Thus, .355 x 1.034=.368.