Factors Affecting the Intake of Dietary Fiber in the United States Diet

Mark C. Senia

Department of Agricultural Economics

Texas A&M University

mcsenia@tamu.edu

Senarath Dharmasena

Department of Agricultural Economics

Texas A&M University

sdharmasena@tamu.edu

Selected Paper prepared for presentation at the Southern Agricultural Economics Association's 2016 Annual Meeting, San Antonio, Texas, February, 6-9 2016

Abstract

The 2010 USDA Dietary Guidelines for Americans recommended that individuals consume around 25 grams of dietary fiber per person per day. Yet despite these recommendations, consumers do not purchase enough foods high in dietary fiber. To investigate the factors behind this behavior, we perform an empirical analysis using Nielsen Homescan data for 2011 with a sample of 62,092 households from across the United States. This research contributes to the literature by simultaneously investigating per capita purchases of products containing fiber: bread, pasta, tortilla, and fresh produce (fruit and vegetables combined). We perform the estimation using a seemingly unrelated regression framework with a Tobit model setting for each equation in order to account for the censored nature of the available data. Preliminary results suggest that individuals that use coupons generally report purchasing more fiber per capita. Those living below 130 percent and below 185 percent of the federal poverty level purchase less fiber per capita from bread and pasta. For these individuals, the fiber purchased from produce is not significantly different from the individuals above the cutoff levels. Those with a higher income or education report more consumption of fiber per capita from produce and less from bread and pasta.

Keywords: Dietary fiber, fiber consumption, Nielson Homescan Panel, USDA Dietary Guidelines

JEL Classification: D12, I18

1. Introduction and Motivation

The 2010 USDA *Dietary Guidelines for Americans* recommend that individuals consume around 25 grams of dietary fiber per person per day. The *Dietary Guidelines for Americans (*USDA, 2005 & 2010) are the main source of dietary recommendations for health professionals and government agencies. One suggestion from these guidelines is to increase the consumption of foods high in fiber. These fiber rich foods include many fruits and vegetables, beans, whole grains, and nuts. Yet despite encouragement from the government, consumers do not purchase enough foods high in dietary fiber. For example, many consumers continue to purchase too few whole grain products (higher in dietary fiber) and too many refined grain products (lower in dietary fiber) than recommended (Volpe and Okrent, 2013).

Dietary fiber provides a range of important health benefits. Consumption of dietary fiber from cereals and fruits is inversely associated with the risk of coronary heart disease (Pereira et al., 2004). Increased intake of dietary fiber may reduce cardiovascular disease, stroke, hypertension, diabetes, obesity, and some gastrointestinal diseases (Anderson et al., 2009; McKeown et al., 2002). There may be an association between adults who eat more whole grains, particularly those higher in dietary fiber, and a lower body weight relative to adults who eat fewer whole grains (USDA, 2010). These health benefits from the dietary fiber make it an important dietary component.

This paper goes beyond a simple total fiber consumption analysis. This paper contributes to the literature by simultaneously investigating three separate purchase categories of products containing fiber: (*i*) bread and pasta purchases (*ii*) tortilla purchases, (*iii*) fresh produce purchases. For each of these categories we calculate the fiber consumption per household member per day. Although the three categories under consideration are all sources of fiber, the amount of each product purchased is driven by a common set of factors unobservable at the individual level to the researcher (food tastes and preferences).

We expect these purchase to be correlated, and therefore, we allow for correlations among corresponding unobservables in the empirical model. By studying the purchases jointly, rather than in isolation from each other, we gain more insight. By jointly estimating the equations for the three purchases while allowing for error terms in them to be correlated, we obtain more efficient estimates of effects of socioeconomic and other variables on the fiber categories. We perform the estimation using a seemingly unrelated regression framework (Zellner, 1962) with a Tobit model setting (Tobin, 1958) for each equation in order to account for the censored nature of the available data.

We perform the empirical analysis using Nielsen Homescan data for 2011. This dataset is well suited to the analysis as information is collected on purchases from participating panelists. The dataset also provides a wealth of socioeconomic data but does have some limitation. This dataset does not provide time spent preparing food and only includes food purchases. The need to account for time is especially important since Aguiar and Hurst (2007) find that households can substitute time for expenditures in food production. One must be careful to differentiation between food that is purchased and food that is consumed.[1] The results of this study can best be interpreted as purchase amount decisions and not consumption amount decisions.[2]

Our main findings can be summarized as follows: Individuals that use coupons report consuming more fiber per capita, those living below 130 percent and 185 percent of the poverty level purchase less fiber per capita from bread and pasta but the fiber consumed from produce

---

[1] The data do not provide information on food that is purchased and given away or food waste.
[2] Though an attempt has been made to ensure the distinction in the paper, it is possible that consumption and purchase may be used interchangeable. The food items purchased in this paper are usually ready to eat and need little preparation time. Thus time inputs are less likely to affect the quality of these goods.

purchase is not significantly different from those above these poverty levels. Those with a higher income or education report a higher consumption of fiber per capita from produce and less from bread and pasta. There also evidence of a cohort effect in that households with an older head of household purchase more fiber per capita. Some regional effects in fiber per capita consumption are also evident but to a less degree than other factors.

The remainder of this paper proceeds as follows. In Section 2, we discuss the existing literature on fiber purchase and consumption. In Section 3, we construct a theoretical model of household fiber consumption to motivate empirical work. In Section 4, we specify the econometric model and outline the estimation methodology. In Section 5, we give a detailed description of the data and the constructed dependent and explanatory variables. In Section 6, we discuss and present the results. In Section 7, we summarize the results and discuss relevance.

2. Literature Review

Current literature dealing solely with consumer dietary fiber demand is somewhat limited. Miguel and Diansheng (2012) use a dynamic Tobit model that allows past purchase occasions to affect current purchase decisions for fiber using the Nielsen Homescan dataset. The authors find that participation in the WIC program, the age and presence of children between thirteen and seventeen, not being Hispanic, and the employment level of the female head do not significantly affect fiber consumption. Also the authors find that the female head's education level has a negative impact on fiber purchases and coupon use has a positive effect.

The effect of nutritional information on nutrient consumption is a popular closely related line of research. Variyam and Blaylock (1996) conduct a survey on the fiber content of food and attitudes toward consumption of foods high in fiber. The authors find that knowledge of nutritional information has an influence on fiber consumption. The major factors affecting fiber intake are household income, meal planner age, smoking status, vegetarian status, race, and ethnicity. Education exerts a sizable intake effect by enhancing the information level. Ollberding, Wolf, and Contento (2011) use the 2005-2006 National Health and Nutrition Examination Survey and find that food label users report a higher fiber consumption than those that do not use food labels. Thus it is likely that in our sample that higher educated individuals will also have higher fiber consumption.

Literature has previously examined the impact of the 1994 Nutrition Labeling and Education Act. Variyam (2008) examined the impact of thirteen nutrients on consumer diets displayed on the consumer nutrition label. The author showed that when consumers use the labels increase their fiber intake of consumers by 0.69 grams per 1000 calories. Using the same data and a different estimation technique, Kim, Nayga and Capps (2000) reported that consumer nutrition label use increase the average daily fiber intake of consumers by 7.51 grams.

Literature has also been focusing on consumer whole grain (a good source of dietary fiber) demand likely due to the USDA making specific quantity recommendations in 2005. Mancino et al. (2008) find that the release of 2005 Dietary Guidelines for Americans increased the availability and sales of whole-grain foods, with a large impact due to reformulation of existing products. Lin and Yen (2008) use the 1994–1996 Continuing Survey of Food Intakes by Individuals to examine how nutrition knowledge and Sociodemographic variables affect the consumption of refined and whole grain products. Mancino and Kuchler (2012) estimate demand for whole grain bread to determine if the release of the 2005 Dietary Guidelines for Americans affected demand for whole grain. They find an increase in demand even after accounting for price changes.

3. Theoretical Model

We first assume weakly separable preferences and two-stage budgeting (Strotz, 1957). The model will be defined beginning in the second stage after the household has chosen the expenditure for the fiber category. Assume household gets utility directly from consumption of the fiber component of the product and not the whole product. We assume that the household decision-maker has well-defined preferences over fiber consumption, a composite commodity comprising the fiber serving for a group of foods. These preferences are represented by the general utility function

$$U = U(BPF, TOF, FPF; \tau), \tag{1}$$

where *BPF* is the commodity "pasta and bread fiber," *TOF* is the commodity "tortilla fiber," *FPF* is the commodity "fresh produce fiber," and τ is a vector of individual characteristics, some of which are observable to the researcher and others are. The utility function *U(·)* in Equation (1) is assumed to be continuously differentiable and concave in the arguments *BPF*, *TOF*, and *FPF*. By specifying that the food commodities enter the utility function separately, we allow for the possibility that different fiber sources are imperfect substitutes and can, in fact, be associated with different utility effects. Also, we assume that the quantities of the fiber commodities cannot be negative. Therefore, $BPF \geq 0$, $TOF \geq 0$, and $FPF \geq 0$.

The household decision-maker faces a conventional linear budget constraint:

$$P_{BPF} \cdot BPF + P_{TOF} \cdot TOF + P_{FPF} \cdot FPF = M, \tag{2}$$

where P_{BPF} is the price of the market commodity BPF, P_{TOF} is the price of the market commodity TOF, P_{FPF} is the price of the market commodity FPF. M is the exogenous expenditure chosen for the category in the first stage of budgeting.

The household decision-maker's problem is to maximize the utility function from Equation (1):

$$max\, U(BPF, TOF, FPF; \tau),$$

subject to the budget constraint (2) and the non-negativity constraints defined earlier. The Lagrangian and Kuhn-Tucker conditions for this problem can be written as

$$\Psi = U(BPF, TOF, FPF; \tau) + \lambda[M - P_{BPF} \cdot BPF + P_{TOF} \cdot TOF + P_{FPF} \cdot FPF] \tag{3}$$

$$BPF: \ \frac{\partial U}{\partial BPF} - \lambda P_{PBF} \leq 0 \qquad BPF \geq 0; \quad \frac{\partial \Psi}{\partial BPF} BPF = 0 \tag{4}$$

$$TOF: \ \frac{\partial U}{\partial TOF} - \lambda P_{TOF} \leq 0 \qquad TOF \geq 0; \quad \frac{\partial \Psi}{\partial TOF} TOF = 0 \tag{5}$$

$$FPF: \ \frac{\partial U}{\partial FPF} - \lambda P_{FPF} \leq 0 \qquad FPF \geq 0; \quad \frac{\partial \Psi}{\partial FPF} FPF = 0 \tag{6}$$

$$\lambda: \quad M = P_{BPF} \cdot BPF + P_{TOF} \cdot TOF + P_{FPF} \cdot FPF \qquad \lambda_1 \geq 0; \quad \frac{\partial \Psi}{\partial \lambda} \lambda = 0 \tag{7}$$

Given that the utility function *U(·)* in Equation (1), a solution for the set of the Kuhn-Tucker conditions (4)–(7) would provide us with an allocation of fiber consumption that maximizes the utility function subject to the specified constraints. The solution to the Kuhn-Tucker problem will be unique due to concavity assumptions placed on the utility function and production functions. The solution would, among other things, specify the optimal amount to consume of each of the three fiber categories in terms of the commodity prices and the individual characteristics affecting the utility:

$$BPF^* = BPF(\, P_{BPF}, P_{TOF}, P_{FPF}, \tau) \tag{8}$$

$$TOF^* = TOF(\, P_{BPF}, P_{TOF}, P_{FPF}, \tau) \tag{9}$$

$$FPF^* = FPF(\, P_{BPF}, P_{TOF}, P_{FPF}, \tau) \tag{10}$$

It should be noted that corner solutions are allowed for in this setting. As an example, it is possible that for some individuals, the optimal consumption of fiber from pasta and bread is zero.

Importantly, notice that Equations (8)–(10) imply that the three fiber consumption equations may each be affected by the same set of unobservable factors (e.g., unobservable food tastes and preferences). Thus, it may be more efficient to estimate these equations together as a system, provided that the estimation procedure properly accounts for the presence of the common unobservables.

4. Empirical Model and Estimation Procedure

As described in the previous section it is possible that the fiber consumption demands can contain a number of zero instances (see Table 1). To account for such observations in the econometric model, we adopt a Tobit model setting (Tobin, 1958; Amemiya, 1984). Since the three fiber consumption equations, Equations (8) – (10), are affected by common unobservable factors, the corresponding error components may be correlated. Thus, we choose to estimate these equations as a system and account for possible correlations by using a Seemingly Unrelated Regression (SUR) modeling framework. Amemiya (1984) and Srivastava and Giles (1987), among others, provide details on the SUR estimation approach. This approach has been used by the author in a situation with similar data limitations (Senia, Jensen, and Zhylyevskyy 2014).

We use notation similar to Huang's (2001)[3]. Let *i* index individuals in the analytical sample, $i = 1, \ldots, n$, where *n* is the sample size (in our case, $n = 62{,}092$). Also, let *j* index the three fiber categories of interest, $j = 1,2,3$, where $j = 1$ is "pasta and bread fiber," $j = 2$ is "tortilla fiber," and $j = 3$ is "fresh produce fiber." Let y_{ij} denote the actual fiber consumption (grams per capita per year) of category *j* as reported by individual *i* for the year. Implicitly, every individual in the analytical sample is assumed to have the opportunity to consume from at least one of the three sources. However, he or she may optimally choose to consume a positive amount of fiber from only one of the activities and not consume from another source (e.g., the fiber consumption from tortillas is zero). Thus, y_{ij} is non-negative, $y_{ij} \geq 0$, and could be left-censored at zero.

Now, we let y_{ij}^* be a continuous latent variable described by the following equation:

$$y_{ij}^* = x_i' \, \beta_j + \varepsilon_{ij}, \tag{11}$$

[3] This is also similar notation from Senia, Jensen, and Zhylyevskyy (2014)

where x_i' is the vector of explanatory variables for individual i β_j is the vector of coefficients (specific to category j) to estimate, and ε_{ij} is the error term. To enable the estimation, we assume that the vector of the error terms for individual i, $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, \varepsilon_{i3})'$, is independent and identically distributed (i.i.d.) across i as a normal random vector, conditional on x_i:

$$\varepsilon_i | x_i \sim i.i.d. \mathcal{N}(0, \Omega), \tag{12}$$

where Ω is a (symmetric) positive definite covariance matrix to estimate.

Because the three fiber consumption categories are likely to be affected by the same set of unobservable factors, we expect that the error terms may, in fact, be correlated with each other. In that case, Ω would be non-diagonal and an efficient estimation of the model parameters could be performed by estimating the three equations jointly as a system. However, if all non-diagonal elements of the matrix Ω were zero, the model could be efficiently estimated equation by equation (Zellner, 1962). Our approach is to estimate the three equations jointly as a system and then test for whether or not Ω is diagonal.

To account for the censored nature of the purchase data, we specify that the observed purchase in activity j, y_{ij}, is related to the latent variable y_{ij}^* as follows:

$$y_{ij} = \begin{cases} y_{ij}^* = x_i' \beta_j + \varepsilon_{ij}, & if \ \ y_{ij}^* > 0 \\ 0, & if \ \ y_{ij}^* \leq 0 \end{cases}, \tag{14}$$

for $j = 1,2,3$. We rewrite this equation more simply as,

$$y_i^* = X_i \beta + \varepsilon_i, \tag{15}$$

where $y_i^* = (y_{1i}^*, y_{2i}^*, y_{3i}^*)'$, $X_i = diag(x_{1i}', x_{2i}', x_{3i}')$, and $\beta = (\beta_1', \beta_2', \beta_3')'$

There are 8 possible combinations of fiber purchase values at their censoring points. The model parameters can be consistently estimated using the maximum likelihood method.

We estimate the model parameters using the SAS QLIM command. This software allows us to implement a SUR estimation for a system of three Tobit equations. To allow interpretation of the results, we calculate and report partial effects associated with the explanatory variables for each of the fiber consumption categories (see Greene, 2012, pp. 848-850):

$$\frac{\partial E[y_{ij} \mid x_i]}{\partial x_i} = \beta_j \ \Phi\left(\frac{\beta_j' \cdot x_i}{\sigma_j}\right), \tag{16}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

The explanatory vector X_i contains a number of important control variables. These include the demographic and economic characteristics from Table 2. Prices are not included at this time since this study is more focused on other factors affecting dietary fiber consumption.

5. Data

Data are obtained from 2011 Nielsen Homescan panel. This consists of a sample of 62,092 households from across the United States. Each participating household is given a scanner to read UPCs from products purchased at stores. Nielsen matches the scanned UPC with products characteristics in their database. The household is also asked to enter quantity, price, and any coupon information about the products.

For a selection of fiber rich products the quantity and demographic characteristics of the household are used. The products selected for study include bread and pasta, tortillas, and fresh produce (fruit and vegetables combined). For each UPC an estimate is made of the fiber content by utilizing UPC keyword search. Then the fiber content for all foods in that category is summed to create the total fiber consumed for the household in that category. This total is then divided by the number of members of the household to create an approximation of fiber consumed per capita.

Table 1 lists summary statistics for these dependent variables. It is important to take notice of the large number of zero observations for the tortilla category. There are a number of households that report unusually large fiber purchases in the categories. This may be due to reporting issues, problems estimating the fiber content of certain foods, or the household purchasing food for members outside of the household (donations to food banks as one possible example).

Figure 1 shows a histogram for the yearly fiber consumption per capita. The black vertical line at 9,125 represents the USDA suggested yearly fiber consumption given 25 grams per day. The white vertical line represents the sample average fiber consumption of 6136 grams per year per person. The majority of our sample are not meeting the USDA guidelines. The USDA (2010, pg. 46) estimates the typical American diet provides 40 percent of needed fiber. Our sample average shows participants meeting 67 percent of the recommendation. This is close to an estimate from 2008 of dietary fiber consumption of 15.9 grams per day (King, Mainous, and Lambourne, 2012).

We begin by including a range of standard demographic characteristics such age, gender, race, and an indicator for Hispanic origin in the empirical model. Table 2 lists summary statistics for these characteristics. The largest segment of individuals fall within the 50 to 64 age group with 45 percent of the sample. The age variable as constructed only takes into account the age of the oldest head of the household. It is assumed that the oldest member is likely to have more influence on purchase decisions. This sample has 5 percent of respondents identifying as being of Hispanic origin. Controlling for Hispanic origin is important because such respondents may have different preferences over the three categories (e.g. more likely to consume tortillas).

The yearly income is included in the explanatory variables. The dataset provides categorical income information. The income variable is constructed as the natural log of the midpoint of the categorical yearly income variable. The average for real income is around $47,100.

The poverty dummy variables indicate whether the respondent's household income is at or below 130% of the federal poverty level and whether the household is at or below 185% of the federal poverty level. The threshold levels here indicate eligibility for participation in public assistance programs such as the SNAP at 130% and below or WIC (Special Supplemental Nutrition Program for Women, Infants, and Children) at 185% and below. By using indicators for eligibility, rather than indicators for actual participation, we avoid potential complications arising from non-random selection into the programs and under-reporting of participation.

Four educational variables denote the highest level of education received. This education variable is constructed to capture the highest level of education completed by either the male or female head of household. It is assumed the individual with the highest education has a large influence in purchase decisions. The sample consists of 53 percent with Bachelor's degree or higher, 29 percent with some college, 17 percent with high school degree, and 1 percent with less than a high school education. This sample is more highly educated than the general U.S. population.[4]

We also construct an indicator for the presence of children in the households. This variable indicates if there is at least one child present in the household. This may be important as the presence for children may change the nutritional mix of food purchased. Parents may focus on purchases healthier food when children are present in the household.

The place of residence dummies use the nine U.S. Census Bureau designated divisions. These are used to control for possible differences in the characteristics of the food environment—including availability of grocery stores and other food outlets, and possible geographical differences in the food tastes and preferences.

6. Results

Table 3 presents parameter estimates for our empirical model. The three food purchase equations are estimated jointly (their parameters are listed in columns (1) through (3) of the

[4] Further research using the Nielsen provided survey weights can make the sample more representative of the U.S. population.

table), and we allow for their error terms to be correlated. As can be seen in the notes to Table 3, the bread and pasta versus tortilla purchase error term correlation is 0.14. An estimate of standard error of this correlation could not be obtained through estimation at this time due to a limitation is SAS. It may not be a stretch to at least assume this is significantly different from zero which would indicate that the covariance matrix of the error terms, Ω, is non-diagonal. Therefore, our choice of the SUR estimation framework is likely appropriate.

The coefficient estimates reported in Table 3 indicate the effects of the explanatory variables on the fiber category purchase being positive. These estimates are not intended to show what impact the variables have on the probability of the purchase being positive. To facilitate the interpretation of the results, we use the coefficients from Table 3 to compute partial effects; see Equation (16) for the formula. The partial effect of an explanatory variable illustrates the influence of a change in this variable on the expected purchase of a category, by accounting for all impacts associated with the change in the variable. We report all partial effects in Table 4 and focus on them when discussing the results below as these more accurately explain the effects of the explanatory variables on fiber purchase. The estimation results reveal statistically significant effects of economic variables on the expected purchase of purchase categories.

Households below the poverty indicators do not significantly differ in their fiber from produce consumption. That these individuals consume a similar amount of fiber from produce to the higher income household is an interesting finding. For these individuals the fiber consumption is lower in the bread and pasta category and for the tortilla fiber category. Another income result is that those with a higher income are associated with getting less fiber from bread and pasta purchases and more fiber from fresh produce purchases. While our results find a negative effect for WIC eligibility on fiber consumption in the bread and pasta category, Miguel and Diansheng (2012) find that participation in the WIC program has no effect on dietary fiber consumption overall.

Some further interesting results arise from the household characteristics. Household that use coupons for produce purchases consume a much larger amount of fiber per capita. There is a similar effect for those that use coupons for a bread or pasta purchase and tortilla purchase. It is possible that household that use coupons are also more likely to accurately report their purchases. An interesting trend is seen with the size of the household. Large households are associated with lower fiber per capita consumption in both the bread and pasta category and the produce category while being associated with higher fiber consumption from the tortilla category. Large households may have more difficulty reporting the large quantity of food that is purchased.

Unexpectedly we find that the presence of children in the household does not affect the fiber per capita consumption. Thus the presence of children does not seem to be associated with healthier eating habits for fiber. This may be caused by only including an indicator for the presence of a child. Further refinement of this specific variable may reveal other effects related to the presences of children. College graduate report 200 grams per capita higher consumption of fiber from produce relative to those with less than a high school degree. Highly educated individuals also consume less fiber from bread and pasta (similar to the income effect).

There also are interesting age effects in the results. Older individuals consume more fiber per capita every category but the tortilla category. The fiber from tortilla consumptions appears to decline with age. It should be noted that the estimated age effects may be accounting for differences in the preferences regarding tortillas across the age cohorts in the sample. In a cross-sectional study such as ours, these possible cohort effects cannot be disentangled from the age effects.

There also appear to be regional differences in fiber consumption. One possible issue with the data is that the fiber consumption from produce is the large significant negative value in most categories relative to the base Pacific category. It may be possible that West Coast households consume more fresh produce due to the proximately to California (a large agricultural production state). It is also possible that this as an issue with the data.[5]

7. Implications and Limitations

This paper constructed a theoretical model of household dietary fiber consumption to motivate an econometric estimation. We specified the econometric model as a system of three food purchase equations estimated jointly. These equation were estimated using Tobit and SUR modeling techniques to account for the censored nature of the fiber consumption and for the correlations among the error terms in the fiber demand equations

A number of interesting finding results from the analysis. Those with a higher income or education report more consumption of fiber from produces and less from bread and pasta. Individuals that use coupons report purchases foods that provide more fiber per capita, those living below 130 percent and 185 percent of the poverty level purchase less fiber per capita from bread and pasta but the fiber from produce purchase is not significantly different from those above these poverty level. There also evidence of a cohort effect in that households with an older head purchase more fiber per capita. Regional effects is fiber purchases are also evident.

While we believe that we accounted for all possible sources of bias in our modeling procedure, limitations of still remain. The expected issues from self-reported data and the restriction to Nielsen households currently prevent generalization to all households.[6] As this is a cross-sectional approach, we cannot be sure that the results truly reflect causation.

---

[5] Those households that report the largest fiber consumption are all mostly located in California.
[6] Einay, Leibtag, and Nevo (2010) have formulated a method to help correct for possible entry errors in the dataset.

Only three categories of fiber were studied and adding additional categories would help capture more fiber consumption. The approach used in this research estimated the fiber content in the food item and this may be the cause of some error. More accurate data that includes the fiber content for each product would improve results. This study is also limited by the lack of data on weighed fresh produce items and this would leave out some fiber consumption. The focus of this paper is food purchased for consumption at home. Fiber consumed away from home would not be captured by this dataset. This may not be a major problem as eating meals away from home is usually associated with less healthy eating (Lin and Guthrie, 2012; Todd, Mancino and Lin, 2010) and this might not change overall fiber totals.

References

Aguiar, M., and E. Hurst. 2005. "Consumption Versus Expenditure." *Journal of Political Economy* 113(5): 919-948.

Amemiya, T. 1984. "Tobit models: A survey." *Journal of Econometrics* 24(1-2):3-61.

Anderson, J.W., P. Baird, R.H. Davis, S. Ferreri, M. Knudtson, A. Koraym, V. Waters, and C.L Williams. 2009 “Health Benefits of Dietary Fiber.” *Nutrition* 67(4): 188-205.

Deaton, A.S., and J. Muellbauer. 1980. “An Almost Ideal Demand System.” *American Economic Review* 70(3): 312-326.

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

Greene, W. H. 2012. *Econometric analysis*, 7th ed. Upper Saddle River, NJ: Prentice Hall.

Huang, H.C. 2001. “Bayesian Analysis of the SUR Tobit Model.” *Applied Economics Letters* 8(9):617-622.

Lin, B., and S.T. Yen. 2008. "Consumer Knowledge, Food Label Use and Grain Consumption in the US." *Applied Economics* 40(4): 437-448.

Lin, B.H., and J. Guthrie. 2012. *Nutritional Quality of Food Prepared at Home and Away From Home, 1977-2008.* Washington DC: Economic Research Service, U.S. Department of Agriculture, Economic Information Bulletin 105, 22 pp. December.

Kim, S.-Y., Nayga, R.M., Jr. and O. Capps. 2000. “The Effect of Food Label Use on Nutrient Intakes: An Endogenous Switching Regression Analysis.” *Journal of Agricultural and Resource Economics* 25(1): 215 - 231.

King, D.E., A.G. Mainous, and C.A Lambourne. 2012. “Trends in Dietary Fiber Intake in the United States.” *Journal of the Academy of Nutrition and Dietetics* 112(5):642-648.

Mancino, L. and F. Kuchler. 2012. "Demand for Whole-Grain Bread Before and After the Release of the Dietary Guidelines.” *Applied Economic Perspectives and Policy* 34(1): 76-101

Mancino, L, F. Kuchler, and E. Leibtag. 2008. “Getting Consumers to Eat More Whole-grains: The Role of Policy, Information, and Manufacturers.” *Food Policy* 33(6): 489–496.

Miguel, H.O. and D. Diansheng. 2012. “Assessing U.S. Household Purchase Dynamics for Dietary Fiber.” Paper presented at AAEA annual meeting, Seattle, Washington, August 12-14.

McKeown, N.M, J.B Meigs, S. Liu, P. WF Wilson, and P.F. Jacques. 2002. "Whole-grain Intake is Favorably Associated with Metabolic Risk Ractors for Type 2 Diabetes and Cardiovascular Disease in the Framingham Offspring Study." *American Journal of Clinical Nutrition* 76(2): 390-398.

Nevo, A. 2000. "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand." *Journal of Economics and Management Strategy* 9(4): 513-548.

Ollberding, N.J., R.L. Wolf and I. Contento. 2011. "Food Label Use and Its Relation to Dietary Intake among US Adults." *Journal of the American Dietetic Assn.* 110(8): 1223-1237.

Pereira, M.A., O'Reilly E, Augustsson K, et al. 2004. "Dietary Fiber and Risk of Coronary Heart Disease: A Pooled Analysis of Cohort Studies." *Arch Intern Med.* 164(4):370-376.

Senia, M. C., Jensen, H. H., & Zhylyevskyy, O. 2014. "Time in Eating and Food Preparation among Single Adults." *Review of Economics of the Household*, forthcoming.

Srivastava, V. K., and D E. Giles. 1987. *Seemingly Unrelated Regression Equations Models: Estimation and Inference*. New York: Dekker.

Strotz, R.H. 1957. "The Empirical Implications of a Utility Tree." *Econometrica*, 25(2):269-280

Tobin, J. 1958. "Estimation of Relationships for Limited Dependent Variables." *Econometrica* 26(1):24-36.

Todd, J.E., L. Mancino, and B.H. Lin. 2010. *The Impact of Food Away From Home on Adult Diet Quality.* Washington DC: Economic Research Service, U.S. Department of Agriculture, Economic Research Report Number 90, February.

USDA. "Dietary Guidelines for Americans." 2005.

USDA. "Dietary Guidelines for Americans." 2010.

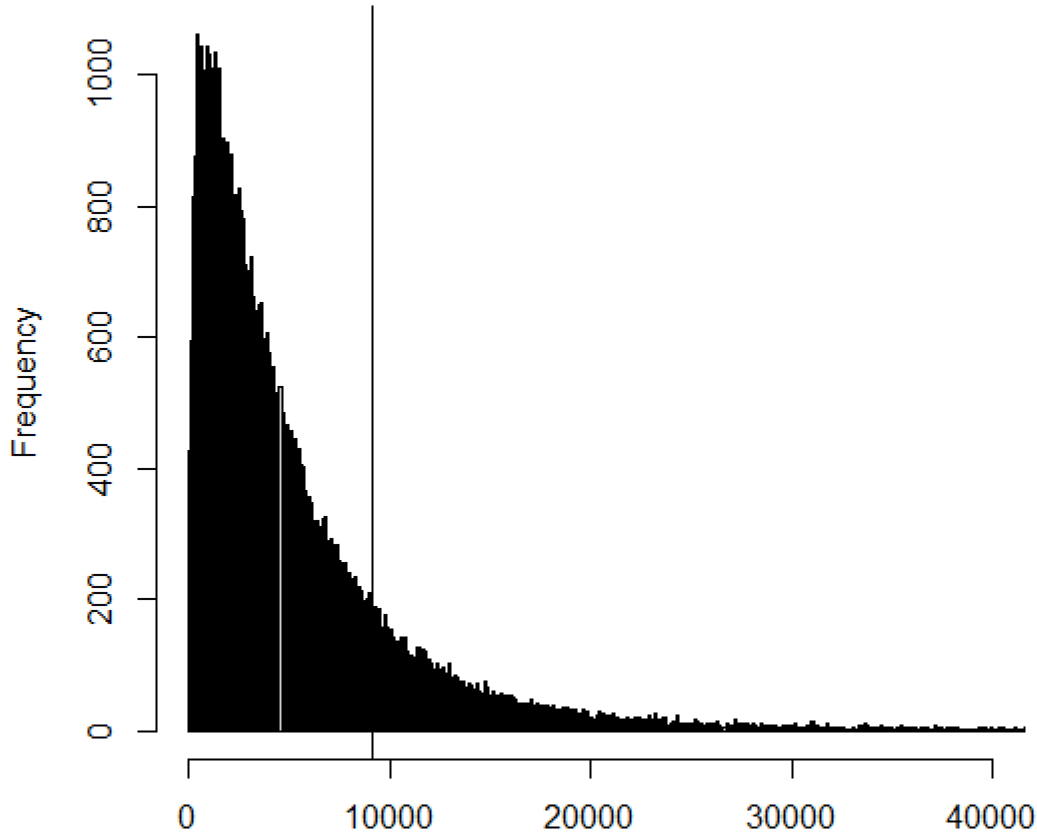
Variyam, J.N. 2008. Do Nutrition Labels Improve Dietary Outcomes? *Health Economics*. 17: 695 – 708.

Variyam, J., Blaylock, J., Smallwood, D., 1996. "A Probit Latent Variable Model of Nutrition Information and Dietary Fiber Intake." *Am. J. Ag. Econ.* 78: 628–639.

Volpe, R., and A. Okrent. 2012. *Assessing the Healthfulness of Consumers' Grocery Purchases*. USDA EIB-102, 34 pp, November.

Zellner, A. 1962. "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias." *Journal of the American Statistical Association* 57(298): 348-368.

Figure 1. Distribution of total year fiber per capita consumption



Notes:
The black vertical line at 9,125 represents the USDA estimated yearly fiber consumption given 25 grams per day. The white vertical line represents the sample average fiber consumption of 6136 grams per year per person.

Table 1. Summary Statistics for fiber consumption categories (yearly grams per capita)

	Bread and Pasta	**Tortilla**	**Fresh Produce**
Sample Mean (std. dev.)	398.3 (364.2)	35.9 (102.8)	5525.5 (25899.5)
Min	0	0	0
Max	10483.6	7500.0	4702592.9
Percent Zero Observation	1.8%	42.8%	3.0%

Notes:
This table lists summary statistics for the three food categories. The percent of zero observations in each category is also listed.

Table 2. Summary Statistics for Explanatory Variables

Variable	Mean	Std. Dev.
Demographic characteristics		
Hispanic origin	0.05	0.22
Age of oldest head of household 65 plus	0.28	0.45
Age of oldest head of household 50-64	0.45	0.50
Age of oldest head of household 40-49	0.19	0.39
Age of oldest head of household 30-39	0.08	0.26
Age of oldest head of household 29 or less	0.01	0.11
Economic characteristics		
Log of family income	10.76	0.67
Income below 130% of poverty line	0.11	0.31
Income below 185% of poverty line	0.25	0.13
Education		
Less than HS degree	0.01	0.11
HS degree	0.17	0.37
Some college	0.29	0.45
Bachelor's or higher degree	0.53	0.50
Family Characteristics		
Used coupon on bread or pasta purchase	0.36	0.48
Used coupon on tortilla purchase	0.06	0.24
Used coupon on produce purchase	0.27	0.44
At least one child present	0.21	0.41
One member in household	0.26	0.44
Two members in household	0.42	0.49
Three members in household	0.14	0.35
Four or more members in household	0.18	0.39
Place of residence		
New England	0.05	0.21
Mid Atlantic	0.13	0.34
East North Central	0.18	0.39
West North Central	0.19	0.39
South Atlantic	0.20	0.40
East South Central	0.06	0.24
West South Central	0.10	0.30
Mountain	0.07	0.26
Pacific	0.12	0.33

Notes:
This table lists summary statistics for the explanatory variables. We report the mean for each characteristic and standard deviations. Categories may not sum to one due to rounding. Except for the log of real family income, all characteristics are indicators. The total number of observations in the analytical sample is 62,092.

Table 3. Estimated Model of the Fiber Purchases

	Bread pasta fiber per capita (1)	Tortilla fiber per capita (2)	Produce fiber per capita (3)
Used coupon on bread or pasta purchase	129.0**(3.0)		
Used coupon on tortilla purchase		99.0**(2.3)	
Used coupon on produce purchase			2193.4**(236.7)
Income below 130% of poverty line	-43.4**(6.5)	-8.1**(2.9)	96.5(424.6)
Income below 185% of poverty line	-23.7**(5.4)	2.6(2.4)	-404.5(311.2)
Two members in household [a]	-53.2**(3.8)	33.3**(1.7)	-844.2**(270.2)
Three members in household	-134.5**(5.5)	32.6**(2.4)	-2191.2**(389.6)
Four or more members in household	-179.7**(6.5)	31.8**(2.8)	-2847.6**(449.5)
Ln Income	-52.7**(4.2)	-4.2*(1.8)	112.0**(36.5)
Hispanic	-13.7*(6.5)	100.5**(2.6)	-329.3(480.6)
New England [b]	58.6**(7.8)	-94.4**(3.5)	-1614.5**(572.3)
Mid Atlantic	48.1**(5.7)	-103.3**(2.5)	-2521.4**(415.0)
East North Central	22.6**(5.3)	-61.0**(2.2)	-2155.9**(386.1)
West North Central	-5.8(6.3)	-47.5**(2.7)	-2477.4**(463.8)
South Atlantic	-15.9**(5.2)	-88.1**(2.2)	-2045.7**(378.0)
East South Central	-17.8*(7.1)	-77.8**(3.1)	-2499.6**(519.5)
West South Central	-22.3**(6.6)	43.2**(2.8)	-491.2(484.3)
Mountain	-7.6(6.6)	-7.8**(2.7)	-1240.4*(485.2)
Child present	-1.3(5.6)	3.6(2.3)	-406.7(407.9)
HS grad [c]	-53.1**(13.6)	-12.2*(6.1)	212.5(194.0)
Some college	-94.8**(13.4)	-7.0(6.0)	132.6(162.9)
College grad or more	-94.8**(13.4)	-12.0*(6.0)	349.1*(152.3)
Age of oldest head of household 65 plus [d]	101.1**(13.6)	-66.1**(5.6)	1650.6**(213.3)
Age of oldest head of household 50-64	90.9**(13.5)	-37.0**(5.5)	1230.5**(213.3)
Age of oldest head of household 40-49	52.9**(13.7)	-22.6**(5.6)	286.5(222.9)
Age of oldest head of household 30-39	27.3(14.3)	-19.6**(5.8)	104.5(307.6)

Notes:
This table presents the estimated coefficients of the seemingly unrelated Tobit model. Standard errors are reported in parentheses. Significance levels are noted as follows: * $p < 0.05$, ** $p < 0.01$. The following present the degree of correlation between error terms: $\rho_{12} = 0.14$; $\rho_{13} = 0.07$; $\rho_{23} = 0.06$;
a Base category is one member in household
b Base category is pacific region
c Base category is less than high school completed
d Base category is age of oldest head of household less than 29

Table 4. Partial Effects of Explanatory Variables

	Bread pasta fiber per capita (1)	Tortilla fiber per capita (2)	Produce fiber per capita (3)
Used coupon on bread or pasta purchase	110.6.		
Used coupon on tortilla purchase		43.9	
Used coupon on produce purchase			1257.9
Income below 130% of poverty line	-37.2	-3.6	**55.3**
Income below 185% of poverty line	-20.3	**-1.1**	**-232.0**
Two members in household [a]	-45.6	14.8	-484.1
Three members in household	-115.3	14.5	-1256.6
Four or more members in household	-154.1	14.1	-1633.0
Ln Income	-45.2	-1.9	64.2
Hispanic	-11.7	44.6	**-188.8**
New England [b]	50.2	-41.9	-925.9
Mid Atlantic	41.2	-45.8	-1446.0
East North Central	19.4	-27.1	-1236.4
West North Central	**-5.0**	-21.1	-1420.8
South Atlantic	-13.6	-39.1	-1173.2
East South Central	-15.3	-34.5	-1433.5
West South Central	-19.1	19.2	**-281.7**
Mountain	**-6.6**	-3.4	-711.4
Child present	**-1.1**	**1.6**	**-233.3**
HS grad [c]	-45.6	-5.4	**121.9**
Some college	-62.5	**-3.1**	**76.0**
College grad or more	-81.3	-5.3	200.2
Age of oldest head of household 65 plus [d]	86.7	-29.3	946.6
Age of oldest head of household 50-64	78.0	-16.4	705.7
Age of oldest head of household 40-49	45.3	-10.0	**164.3**
Age of oldest head of household 30-39	**23.4**	-8.7	**59.9**

Notes:
This table presents average partial effects of explanatory variables in Table 3 on the conditional purchase of food categories. These are calculated using equation (16). Refer to Table 3 for significance levels as the marginal effects will have the same significance as their respective coefficients. For quick reference those in bold represent those that failed to reached the 5% significance level.
a Base category is one member in household
b Base category is pacific region
c Base category is less than high school completed
d Base category is age of oldest head of household less than 29